QUATERNION FRAMEWORK FOR LOW-LIGHT IMAGE ENHANCEMENT

A THESIS REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
ARTIFICIAL INTELLIGENCE

Submitted by

SUDEEP KUMAR ACHARJEE
(23/AFI/08)

Under the supervision of

Prof. ANIL SINGH PARIHAR

COMPUTER SCIENCE & ENGINEERING DEPARTMENT
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

MAY, 2025

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

<u>CANDIDATE'S DECLARATION</u>

I, **SUDEEP KUMAR ACHARJEE** (Roll No – 23/AFI/08) student of M.Tech in Dept. of Computer Science & Engineering, hereby declare that the thesis titled "**Quaternion Framework for Low-Light Image Enhancement**" which is submitted by me to the Computer Science & Engineering Department, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of degree of Master of Technology in the period from August 2023 to May 2025 under the supervision of Dr. Anil Singh Parihar, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi   Sudeep Kumar Acharjee

Date: 10.06.25   M.Tech - AFI (23/AFI/08)

COMPUTER SCIENCE & ENGINEERING DEPARTMENT
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the thesis titled "**Quaternion Framework for Low-Light Image Enhancement**" which is submitted by **Sudeep Kumar Acharjee**(Roll No – 23/AFI/08), Dept. of Computer Science & Engineering Department ,Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of **Master of Technology in Artificial Intelligence**, is a record of the research work carried out by the students under my supervision during the academic session 2023-25. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

EXTERNAL EXAMINER   **SUPERVISOR**

_____   Prof. Anil Singh Parihar

_____   Computer Science & Engineering

_____   Delhi Technological University

Delhi - 1100____   Delhi - 110042

Place: Delhi

Date: 31.05.2025

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I wish to express my sincerest gratitude to **Prof. Anil Singh Parihar** for his continuous guidance and mentorship that he provided us during the project. He showed us the path to achieve our targets by explaining all the tasks to be done and explained to us the importance of this project as well as its industrial relevance. He was always ready to help and clear doubts regarding any hurdles in this research. Without his constant support and motivation, this project would not have been successful.

I would also like to extend my heartfelt thanks to **Dr. Kavinder Singh, Assistant Professor**, for his valuable insights, constructive feedback, and constant encouragement throughout the course of this research. His deep knowledge and thoughtful suggestions played a crucial role in refining the technical depth and direction of this work. His willingness to support me at every stage is sincerely appreciated.

Place: Delhi

Sudeep Kumar Acharjee

Date: 31.05.2025

(23/AFI/08)

Abstract

Images captured at night affected by various degradation such as color distortion, low contrast, and noise. Many existing methods improve low-light images may sometimes amplify noise, causes color distortion, and lack finer details. The existing methods require larger number of parameters, which limits the adoption of these method in vision-based applications. In this thesis, we proposed a QLight-Net method to achieve a better enhancement with a comparably lower number of parameters. We proposed a novel method based on quaternion number representation utilizing a quaternion-based convolutional layer and introducing quaternion based depth-wise convolution, and quaternion cross attention to develop the two-branch architecture to enhance the input low-light images in this thesis. The proposed dual branch model leverages gradient branch to extract color-aware gradient features from the training images, and it uses color branch to extract gradient-aware color features. The dual branch architecture inspired from the U-Net architecture incorporates residual connections in the respective encoder and decoder blocks. The branch leverages the knowledge and learning from the opposite branch throughout the model by sharing cross information sharing methodology after each layer of encoder and decoder. The proposed model achieve LPIPS score of 0.048, which surpasses the previous best results and parameter count of 0.7 million with 0.8804 and 28.05 scores of SSIM and PSNR, respectively in LOLv1 dataset. Our approach attains a balance between the computational efficiency and high-quality enhancement, providing improved results without the computational cost of previous state-of-the-art models while trained on LOLv-1 dataset. The results are compared with 18 existing low-light image enhancement models each contributing unique methods to enhancement and shows that it outperforms many quantitatively.

Contents

List of Tables

List of Figures

List of Symbols

$\hat{q}$	Quaternion
$\uparrow$	Measure that shows higher value is better
$\downarrow$	Measure that shows lower value is better
$\hat{I}$	Enhanced output from model
w/o	Without(exclusion of something)
w	With(inclusion of something)

Chapter 1

INTRODUCTION

The rapid growth of digital imaging in diverse fields, including intelligent systems, has driven a need for image enhancement. From military surveillance to underwater photography, digital images are often captured in challenging low-light conditions, necessitating reliable enhancement methods. Photographs taken outdoor in low-light environment may suffer from low-light, lens flare or extra noise. Images shot in low-light conditions may affect the decision-making capabilities of autonomous vehicles. Low-Light image enhancement focuses mainly on improving the brightness while maintaining the color consistency and reducing noise, and impact of color bias [8].

In the past decades, an extensive development in the domain of low-light image enhancement task proposed variety of approaches from traditional methods in the pre-deep learning era to using of diffusion models in recent years. One of the earliest methods [9] that introduces histogram equalization based method that proposed iterative histogram modification that improves the contrast of digital images. Early works laid the groundwork for modern image enhancement techniques that use gamma correction for brightness and retinex theory for enhancing the images that suffer from low lighting environment. Another work [10] created a foundational ground by introducing the Retinex theory, explaining how human vision perceives color under varying lighting conditions by decomposing the image into illumination and reflectence components. Later, Edwin Land expands on the Retinex theory [1] by emphasizing how the visual system maintains color consistency in different lighting conditions and to utilize it on the image enhancement tasks. Since the introduction of Convolutional Neural Networks (CNN), many CNN-based method for low-light image enhancement task have been proposed, and laid the groundwork for using CNN's for image enhancement tasks. In LLNet [11], it introduces one of the intial CNN-based models for low-light enhancement, using a deep auto-encoder to improve brightness and reduce noise in the low-light image. RetinexNet [1] combines the concept of retinex theory with CNN to decompose the image into illumination and reflectance. The works focus on decomposing and adjusting the illumination component. In parallel there's been rapid development in unsupervised methods. Jiang et al. [12] used GAN technique that helps in enhancing extremely low-light images without paired training data while preserve information.

While these approaches achieved significant performance improvement, they also present certain limitations that motivate this research. Most existing deep learning-based methods rely on large, complex architectures with millions of parameters. Such models are computationally expensive, which limits their deployment on edge devices such as autonomous vehicles and mobile applications. Although some lightweight models are proposed [13], many of them compromise on enhancement quality or fail to retain structural and color details. Furthermore, current approaches typically operate in real-valued space, resulting

1

in a higher number of parameters. These models also struggle to explicitly model the interaction between different types of features, which is crucial for producing perceptually realistic results in low-light conditions. There is a pressing need for a method that enhances brightness and contrast but ensures adequate computational efficiency too.

To address the challenges, we proposed **QLight-Net**, a lightweight model for low-light image enhancement task. The design of QLight-Net is motivated by the representational efficiency of Quaternion Convolutional Neural Networks[14] (QCNNs), which leverage quaternion algebra to encode and process multi-channel image features in a compact form. By employing the Hamilton product, QCNNs reduce the number of learnable parameters while retaining expressive power. However, we further optimize this architecture by introducing depth-wise quaternion convolution, which processes features channel-wise and significantly reduces both training and inference cost. Additionally, low-light enhancement not only requires brightness correction but also demands preservation of structural details and color consistency. To address this, we design a dual-branch architecture, where one branch is developed to capture gradient-based features and the other for color information. The outputs of both branches are fused to produce enhanced images that are both perceptually pleasing and structurally coherent.

The key contribution of our QLight-Net are as follows:

1. We introduced a novel Depth-wise Quaternion Convolutional Layer that reduces the number of parameter and computational cost while maintaining the performance. The module builds on and simplifies traditional quaternion convolutional layers, enabling efficient deep learning for low-light enhancement task.

2. We propose a Quaternion Cross-Attention Encoder-Decoder architecture for low-light enhancement. This design enables robust feature extraction by leveraging the representational power of quaternion algebra and attention mechanisms.

3. We design a two-branch architecture to emphasize both structural and color information. The Gradient Branch captures edge and intensity cues, while the Color Branch focuses on chromatic fidelity. These branches are fused for optimal enhancement results.

4. Our model outperforms several SOTA approaches on both quantitative and qualitative, benchmarks while using significantly fewer parameters. This demonstrates an effective trade-off between the performance and computational efficiency.

Chapter 2

LITERATURE REVIEW

In this section, we discussed various papers focusing on advancements in the field of low-light image enhancement task. The domain witnessed significant progress over the years, with the significant growth of models being deep learning-based. However, some notable works [15, 16], explores some advance techniques as well like diffusion models (used in various domains [17, 18, 19]) and attention mechanism. Initially, many studies relied on Retinex theory-based methods for enhancement. LIME [20] proposed an illumination estimation technique inspired by Retinex theory. It estimates illumination using the maximum of RGB channels and adjusts brightness and noise through the reflectance layer to produce the final enhanced image. Several methods also proposed hybrid techniques by combining Retinex theory with other mechanisms. Ma et al. [21] introduced a decomposition network that generates high-quality illumination and reflectance maps by leveraging attention mechanisms for improved results. Liu et al. [22] presented retinex inspired unrolling model that dynamically adjusts the illumination component. Several models have introduced unique and advanced methods, employing diverse approaches to tackle the challenges of low-light image enhancement effectively. Restormer [23] utilized a transformer-based methods that demonstrated exceptional performance in enhancing high-resolution images. However, the practical applicability of these models remains limited due to their high computational demands. While most models rely on supervised learning, there has been some progress in semi-supervised and unsupervised techniques as well. Yang et al. introduced a semi-supervised model [24] that aims to balance the fidelity and perceptual quality during the learning process. Further, EnlightenGAN [25] employed an unsupervised learning method. Unsupervised techniques hold significant potential and useful in real-world applications, but developing effective models using this approach is highly challenging, as the methods like EnlightenGAN struggles to preserve fine details, particularly in extreme low-light conditions that highlights the complexities of such techniques. Zhou et al. proposed LEDNet [26], which attempts to address the low-light and motion blur issue simultaneously. However, achieving a balance between these two tasks becomes very difficult in extreme cases. Additionally, images that vary in lighting conditions across different regions impact the performance of some models. Xu et al. introduced the SNR-Aware approach [27], which considers the noise levels to improve image quality. However, this method struggles to balance and differentiate effectively under varying lighting conditions. Probabilistic deep learning techniques are also explored for this task. LLFlow [28] uses a normalizing flow-based probabilistic generative modeling technique that tries to establish a bijective-mapping between the low-light and enhanced images by leveraging the distribution of normal-light images during traning. While it produces high-quality enhanced images, the significant weight of the model limits its practical applications, posing a major drawback. Wang et al. proposed LLFormer [29],

a transformer-based model that excels at enhancing fine details in an image. However, it lacks in computationally efficiency for practical use. PairLIE [30] was introduced in 2023 as a lightweight model that learns from paired data. While effective, it faces adaptability challenges in handling complex regions within image. Balancing between the model weight or parameter count and enhancement results remains a critical and challenging trade-off in this domain.

The domain of low-light image enhancement has developed significantly through the application of deep learning techniques. Several methods integrate Retinex theory into deep learning frameworks to enhance images effectively. Wei et al. [1] proposed a way that utilizes and trained on LOL dataset and decomposes each input images into illumination and reflectance components, facilitating low-light image enhancement. Another approach introduces a dual-information based network that trained spatial and channel attention mechanism that produces competitive performance in [31]. While high-quality enhancement is important factor, number of parameters in a model equally critical. The approach proposed in Zero-Reference Deep Curve Estimation [32] is lightweight architecture that estimates pixel-wise curves without requiring paired input and ground truth data for training. Further, Li et al. [33] proposed a parametric Retinex model utilizes gradient based regularization, which enhances illuminance component and reduces noise in low-light images. Liu et al. [22] unrolls search algorithm motivated by Retinex model that introduces optimization-based technique for image enhancement. Two network based architectures are extensively getting used in recent advancement. In BLNet architecture [34], two UNets based method is proposed with different role to learn, with one dedicated in noise removal and another on color restoration, which makes it robust and fast enhancement technique. Additionally, EEMEFN [35] merges two different models that is for edge restoration and color preservation that enhances superior quality enhanced images. All the mentioned approaches shows the diversity of methods in the low-light image enhancement technique, each methods addresses a unique challenges such as de-noising, color restoration, and lightweight architectures, which remain critical for real-world applications.

Zhu et al. [36] introduced a novel approach that applies ghost imaging for multi-illumination estimation. It reconceptualizes the LLIE task as a re-imaging process using structured light estimation and differential ghost imaging. The method integrates light modulation, gradient-guided denoising, and color adaptation to tackle over/underexposure. Abbass et al. [37] developed transformer-based YUV network, YUVAtten-Net for low-light enhancement. By leveraging the luminance-chrominance separation inherent in the YUV color space, it designed a triplet-attention-based denoising and fusion framework that achieves enhanced perceptual quality with reduced computational requirement. Xu et al. [7] proposed BiFormer, a hybrid CNN-Transformer framework that harmonizes local detail preservation and global consistency. The model develops collaborative perception which integrates a local-aware convolutional attention module and a global-aware recursive transformer. Ma et al. [38] addressed the challenge of low-light in the wild, where diverse illumination zones coexist within a single image. RLLIE leverages Explicit Domain Supervision (EDS) to perform unsupervised segmentation and applies region-specific enhancement strategies. This region-based paradigm marks a shift from monolithic enhancement models toward more granular and context-aware frameworks.

2.1 Background: Quaternion Number System

Quaternion number system introduced the four-dimensional concept of extended complex number, which provides the framework to represent 3D rotation and information of four channel in a single channel. Since the introduction, the advantage of quaternion has been leveraged in many concepts. The introduction of Gabor Filters[39] for texture segmentation applied quaternion mathematics in image processing by capturing multi-channel information more effectively. Uses of quaternion concepts in fourier transform[40] for color image processing to jointly process RGB channel while preserve color relationships. And, applying the quaternion concepts in wavelet transform[41] for color image enhancement. From the development of QCNN(Quaternion-Convolutional Neural Networks)[14] and[42] that utilizes the quaternion algebra for RGB channel while enabling the network to model inter-channel correlations more effectively than traditional real-valued CNNs for image classification and forensics tasks.

Quaternion algebra extends the complex numbers to four dimensions, with one real valued and three imaginary components.

$$q = r + x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} \tag{2.1}$$

Here r is the real-valued, and $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ being the imaginary part and x, y, z are real numbers.

2.1.1 Quaternion Properties

Key properties of quaternion number space:

Non-Commutative Multiplication: Quaternion multiplication is not commutative.

$$q_1 \cdot q_2 \neq q_2 \cdot q_1 \tag{2.2}$$

Imaginary Unit Multiplication Rules: The following rules define the multiplication of the imaginary units:

$$\hat{i}^2 = \hat{j}^2 = \hat{k}^2 = \hat{i}\hat{j}\hat{k} = -1$$
$$\hat{i}\hat{j} = \hat{k} \quad ; \quad \hat{j}\hat{i} = -\hat{k}$$
$$\hat{j}\hat{k} = \hat{i} \quad ; \quad \hat{k}\hat{j} = -\hat{i}$$
$$\hat{k}\hat{i} = \hat{j} \quad ; \quad \hat{i}\hat{k} = -\hat{j}$$

Conjugation: The conjugate of a quaternion $q = r + x\hat{i} + y\hat{j} + z\hat{k}$ is:

$$q^* = a - x\hat{i} - y\hat{j} - z\hat{k} \tag{2.3}$$

Norm: The norm of a quaternion is given by:

$$\|q\| = \sqrt{r^2 + x^2 + y^2 + z^2} \tag{2.4}$$

2.1.2 Multiplication(Hamilton Product):

In quaternion multiplication operation was introduced by *William Rowan Hamilton*[43].

$$q_1 = r_1 + x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}}$$

$$q_2 = r_2 + x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}} + z_2\hat{\mathbf{k}}$$

It combines the components according to specific algebraic rules as follows-

$$
\begin{aligned}
q_1 \times q_2 = {} & (r_1r_2 - x_1x_2 - y_1y_2 - z_1z_2) \\
& + (r_1x_2 + x_1r_2 + y_1z_2 - z_1y_2)\hat{\mathbf{i}} \\
& + (r_1y_2 - x_1z_2 + y_1r_2 + z_1x_2)\hat{\mathbf{j}} \\
& + (r_1z_2 + x_1y_2 - y_1x_2 + z_1r_2)\hat{\mathbf{k}}
\end{aligned}
\tag{2.5}
$$

The Hamilton product can be represented as matrix-vector multiplication which can be represented as -

$$
q_1 \times q_2 =
\begin{bmatrix}
r_1 & -x_1 & -y_1 & -z_1 \\
x_1 & r_1 & -z_1 & y_1 \\
y_1 & z_1 & r_1 & -x_1 \\
z_1 & -y_1 & x_1 & r_1
\end{bmatrix}
\begin{bmatrix}
r_2 \\
x_2 \\
y_2 \\
z_2
\end{bmatrix}
\tag{2.6}
$$

The quaternion q_1 is transformed into a quaternion matrix form and quaternion q_2 is represented in vector form.

2.1.3 Quaternion-Softmax

The Softmax function is commonly used in machine learning to transform a vector into a probability distribution.
Let a vector of quaternion inputs $Z = [q_1, q_2, .., q_n]$, where $q_i = r_i + x_i\hat{\mathbf{i}} + y_i\hat{\mathbf{j}} + z_i\hat{\mathbf{k}}$.

$$
Q_{SM}(q_i) = \frac{e^{\|q_i\|}}{\sum_{j=1}^{n} e^{\|q_j\|}}, \quad \text{for } i = 1, 2, \ldots, n
\tag{2.7}
$$

Where $\|q_i\|$ is norm of quaternion q_i.

2.2 Color Space Transform

The color space transform proposed in [44], addresses the limitations of RGB color space under low-light conditions by converting the input RGB image into a derived color space based on the HSV model. In low-light images, brightness(value V) is significantly reduced, while color information(hue H, and saturation S) may still be preserved. Transforming the image allows independent processing of hue, saturation and value, preventing color distortion during enhancement. The brightness(value V) is computed as the maximum intensity across the channels along red, green and blue channels. Hue(H), representing the color tone, depends on the dominant color channel and is adjusted with specific offset for red,green and blue. For gray-scale pixels, the hue is set to zero. Saturation(S), which indicates the color intensity, is calculated as the normalized difference between the maximum and minimum intensities, and it is set to zero when brightness is zero.

To handle brightness sensitivity under low-light conditions, a trigonometric sine function simulates non-linear sensitivity to brightness, influenced by a learnable parameter k, as shown in 2.8:

$$C_{\text{sensitive}} = (\sin(V \times 0.5 \times \pi) + \epsilon)^k \tag{2.8}$$

Here, $\sin(V \times 0.5 \times \pi)$ outputs values between [-1, 1], and ϵ to avoids zero values. Raising the result to the power k adjusts the sensitivity to brightness.

The hue is transformed into angular components:

$$
\begin{aligned}
c_{x-axis} &= \cos(2\pi \times H) \\
c_{y-axis} &= \sin(2\pi \times H)
\end{aligned}
\tag{2.9}
$$

This represents the hue directionally in a 2D plane.

The final output channels are derived as:

$$
\begin{aligned}
C1 &= C_{\text{sensitive}} \cdot S \cdot c_{x-axis} \\
C2 &= C_{\text{sensitive}} \cdot S \cdot c_{y-axis} \\
C3 &= V
\end{aligned}
\tag{2.10}
$$

Here, C_1 and C_2 encode color information modulated by brightness sensitivity, while C_3 directly retains the brightness(V). This three-channel output ensures enhanced low-light images while maintaining accurate color representation.

2.3   Gradient Feature Map

In the paper[45], it uses the concept to detect vague edges for the transportation surveillance under low-light environments which badly suffers from low brightness in the environment. It tries to find the segments where there is sudden intensity shift at the pixel level. Images shot in low-lighting conditions affected from reduced visibility because of low brightness and sensor limitations, To address this, edges and boundaries are first identified, as they are crucial for object detection. The RGB image undergoes Gaussian smoothing(kernel size 3) to reduce noise and soften edges, followed by scaling and clipping to ensure proper intensity range. The Laplacian operator then preserves edge details by applying second-order derivatives. Finally, a Gradient Feature Map is generated by normalizing pixel values to an 8-bit range, enhancing the model's ability to differentiate noise from true edges effectively.

Chapter 3

METHODOLOGY

In this section, it contains detailed discussion of proposed architecture **QLight-Net**. The proposed QLight-Net is a lightweight model with fewer parameters as it focuses on parameter sharing. We developed a new depth-wise quaternion convolution layer to effectively process the spatial features within a set of channels, enabling fine-grained feature extraction while preserving computational efficiency. The framework of proposed QLight-Net illustrated in Fig. 3.1. QLight-Net consists of two distinct branches; Gradient Branch (GB) and Color Branch (CB). The input image is simultaneously processed to obtain log-based gradient and HVI color space [44]. The GB processes log-based gradient combined with the intensity channel of HVI color space. The CB processes HVI color space concatenated with the log-based gradient. The GB and CB are U-Net inspired architectures with cross information sharing structure consists of Color-aware Gradient Encoder-Decoder(CGE-D) and Gradient-aware Color Encoder-Decoder (GCE-D), respectively.
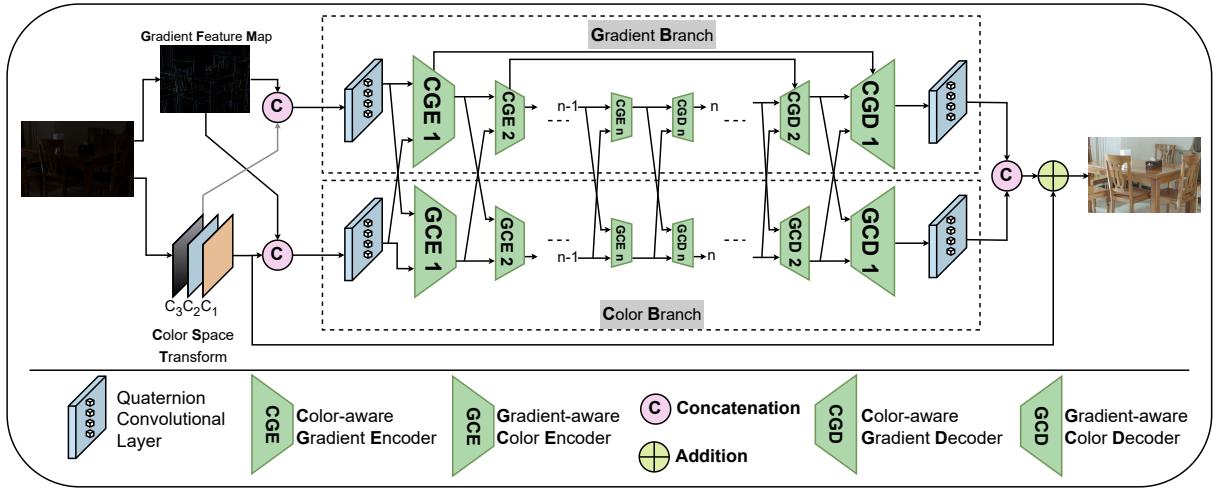
Figure 3.1: QLight-Net: Overview of the proposed low-light image enhancement architecture. The architecture features two branches in parallel: the Gradient Branch (GB) and the Color Branch (CB). GB enhances edge details using log-based gradients, as demonstrated in the result shown above, where zooming in reveals edges extracted from a low-light image. CST restores color information by transforming the input into a 3-channel modified HSV color space. Attention mechanisms are applied across n-stages to progressively refine image features. The outputs from both branches are fused using quaternion convolutions and residual connections, followed by a addition operation to ensure accurate color restoration and preserve image details.

The architectural details of CGE-D & GCE-D are discussed in 3.2 and 3.3, respectively.

The detailed discussion of the proposed structure are presented in 3.4.

3.1 Depth-wise Quaternion Convolutional Layer (DQC)

In traditional depth-wise convolution [46], each input channel is convolved separately with its own filter, effectively isolating the spatial features of each channel but ignores the inter-channel correlations. To address this limitation, we propose Depth-Wise Quaternion Convolution (DQC), a new approach that combines the efficiency of depth-wise operations with the quaternion algebraic representation. DQC processes four input channels together, treating them as components of a quaternion, which inherently models inter-channel relationships. By leveraging quaternion algebra, the proposed DQC captures both intra-channel spatial features and inter-channel dependencies, resulting in enhanced feature representation and processing.

The DQC performs quaternion convolution independently on every set of 4 input channels representing a single quaternion layer. Each group consists of 4 channels, which are processed without filter sharing. To match the output dimensions and integrate the extracted features, we further introduce a point-wise quaternion convolution, as represented in Fig. 3.2. Given a feature map $\mathbf{X} \in \mathbb{R}^{H \times W \times C_{in}}$ (height H, width W, and C_{in} channels), where C_{in} is divisible by 4, we group every 4 channels into a quaternion-valued features representing single quaternion layer, represented as:

$$\mathbf{X}(i,j,g) = x_r(i,j,g) + x_i(i,j,g)\hat{\mathbf{i}} + x_j(i,j,g)\hat{\mathbf{j}} + x_k(i,j,g)\hat{\mathbf{k}} \tag{3.1}$$

where $g = 1,...,G$ is the quaternion group index, $G = C_{in}/4$, and $x_r, x_i, x_j, x_k \in \mathbb{R}^{H \times W \times G}$ are real-valued quaternion components.
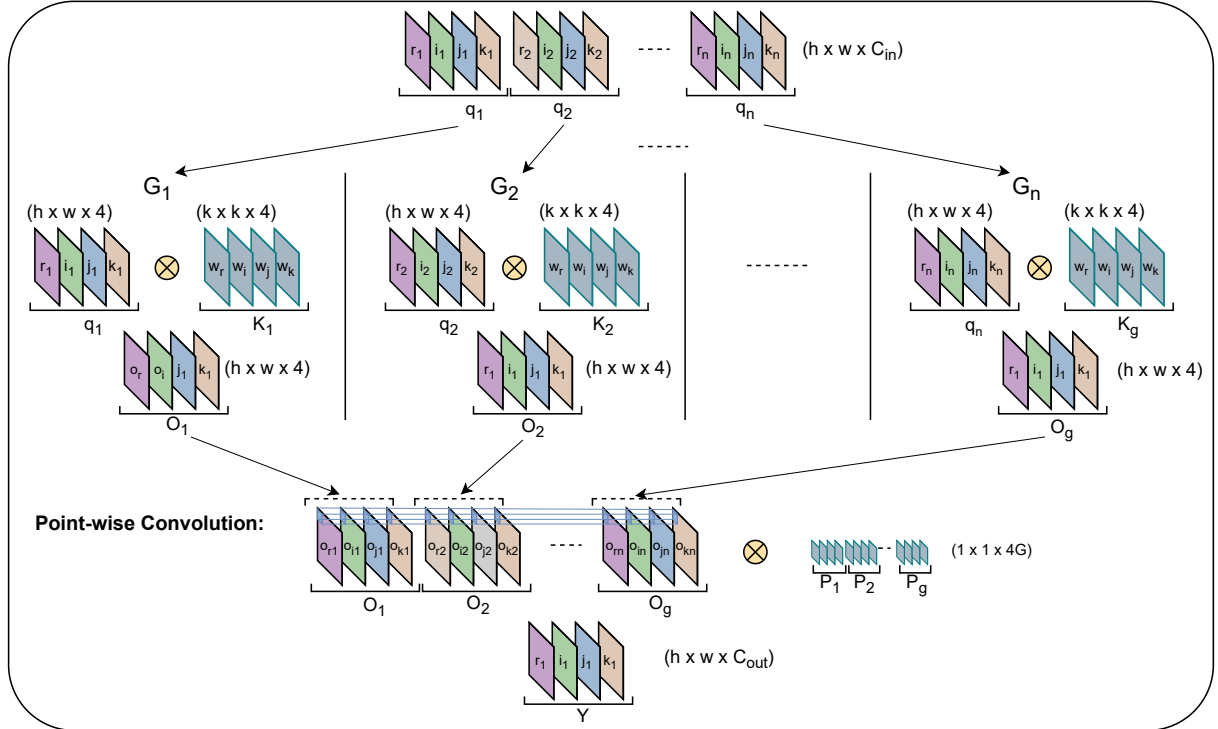
Figure 3.2: Framework of our Depth-wise Quaternion Convolution (DQC).

The Depth-wise Quaternion Convolution processes each quaternion group independently, applying a quaternion kernel of size $k{\times}k{\times}4$ without parameter sharing across groups, where 4-depth corresponds to the quaternion components. The kernel for a single quaternion group is defined as:

$$\mathbf{K}(m, n, 4) = w_r(m, n, 4) + w_i(m, n, 4)\hat{\mathbf{i}} + w_j(m, n, 4)\hat{\mathbf{j}} + w_k(m, n, 4)\hat{\mathbf{k}} \qquad (3.2)$$

where $m, n \in \{0, ..., k - 1\}$, and w_r, w_i, w_j, w_k are the real-valued components of the quaternion filter $\mathbf{K}$. The quaternion convolution operation between the input $\mathbf{X}$ and the kernel $\mathbf{K}$ for a single group g is:

$$(\mathbf{X} * \mathbf{K})_g(i, j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} \mathbf{X}_g(i + m, j + n) \otimes \mathbf{K}_g(m, n) \qquad (3.3)$$

where $\otimes$ denotes hamilton product. Each quaternion group is convolved independently, ensuring inter-channel dependencies within each quaternion group are maintained.

After depth-wise quaternion convolution, the output tensor has the same number of groups G, where each group has 4 channels. To merge these features and match the desired output channels C_{out}, we apply a point-wise quaternion convolution with a $1{\times}1{\times}C_{out}$ kernel. The point-wise quaternion kernel $\mathbf{P}$ for a group is expressed as:

$$\mathbf{Y}_g(i, j) = \mathbf{O}_g(i, j) \otimes \mathbf{P}_g(i, j) \qquad (3.4)$$

where $\mathbf{O}_g(i, j)$ is the output of the depth-wise quaternion convolution for group g. The final output $\mathbf{Y} \in \mathbb{R}^{H \times W \times C_{out}}$ is obtained by convolution operations and appropriately reshaping.

The proposed DQC inherently models cross-channel dependencies through its quaternion structure, ensuring that the relationships between channels, essential for tasks requiring color and texture consistency, are effectively captured. The capability is particularly beneficial from low-lighting scenarios, where subtle inter-channel variations play a critical role. Moreover, the representation of DQC is substantially enhanced by its ability to encode both spatial (edge and structure) and spectral (color and intensity) information within a single quaternion framework. This richer feature representation allows for better handling of complex low-light scenarios that traditional DQC often struggles to address. By preserving the relationships between channels, DQC ensures alignment and consistency across the feature space, which is pivotal in low-light conditions where phase shifts are common due to uneven lighting or noise. For instance, when the dominance of one channel shifts relative to others, DQC effectively models and balances this shift, preserving the integrity of the reconstructed image.

3.2 Color-aware Gradient Encoder (CGE) and Decoder (CGD) Block

The Gradient Branch in our model contains Color-aware Gradient Encoder and Decoder blocks to process the log-based gradient features. The CGE and CGD blocks leverages the new quaternion cross attention to process the gradient features which are aware with the color features. The low-light images often suffer from inadequate edge definition and diminished contrast, therefore the encoders and decoders benefit from cross-branch feature sharing to compensate for the deficiencies.
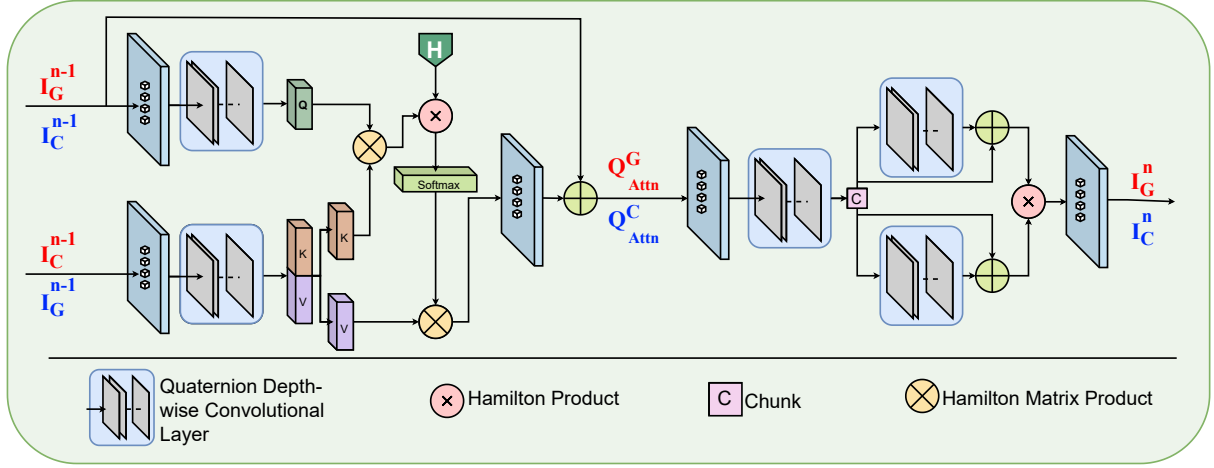
Figure 3.3: Illustrates the architectures of the Color-aware Gradient Encoder-Decoder(CGE-D) and Gradient-aware Color Encoder-Decoder(GCE-D) modules. This module in both the branches take two inputs: one from the GB and another from CB. To distinguish the inputs and output from the module, notations marked in *Red* and *Blue* are colors represents Color-aware Gradient Encoder-Decoder and Gradient-aware Color Encoder-Decoder, respectively. In GB, I_G^{n-1} and I_C^{n-1} are the inputs from the $n-1^{th}$ module of CGE-D from GB and GCE-D from CB, respectively.And in CB, I_C^{n-1} and I_G^{n-1} are the inputs from the $n-1^{th}$ module of GCE-D from CB and CGE-D from GB, respectively. Q_{Attn}^G and Q_{Attn}^C are the attention computed in both the branch in GB and CB respectively, which are being pass forwarded. In GB before producing the final output, it performs an addition operation with Q_{Attn}^G as a skip connection and produces I_G^n, while in CB it directly outputs I_C^n without performing any extra operation.

In the GB, the CGE and CGD blocks operate on the gradient features derived from the image. These gradient features are then processed through quaternion convolutional layer followed by a DQC. These steps effectively refine the spatial and channel-wise dependencies of the feature map, enhancing the gradient features to be more localized and contextually aware. The focus is that the GB should be color-aware despite being primarily concerned with gradient features. To achieve this, the features from the CB are shared with the GB to ensure that the gradient-based attention mechanism in the GB is not only sensitive to the edges but also incorporates color-related components.

The attention mechanism in the CGE and CGD block is based on the dual-input strategy: one input (I_g^{inp}) comes from the Gradient Branch, while the other input (I_c^{inp}) is obtained from the Color Branch. This approach ensures that the attention mechanism considers the importance of both gradient and color features effectively. Specifically, the quaternion query (C_q) is calculated from the color feature map I_g^x of GB, and the quaternion key (G_k) and value (G_V) are calculated from the gradient features I_c^x of the CB, which can be represented as:

$$G_q = \mathbb{Q}_{DQC}[\mathbb{Q}_{conv}(I_g^x)] \tag{3.5}$$

$$C_k, C_v = \mathbb{Q}_{DQC}(\mathbb{Q}_{conv}(I_c^x)) \tag{3.6}$$

where I_g^x and I_c^x represents the xth features from the GB and CB, respectively. The quaternion convolution ($\mathbb{Q}_{conv}$) and depth-wise quaternion convolution ($\mathbb{Q}_{DQC}$) are used to process these inputs, ensuring that both spatial and channel-wise dependencies are

captured efficiently. The input I_g^x is first passed through quaternion convolution layer, which doubles the number of feature maps. These features maps are then evenly split into two parts C_k and C_v after the depth-wise quaternion convolution layer, ensuring that the G_q, C_k and C_v have the same dimensions. This step aims to reduce computational complexity while preserving relevant feature information for both branches.

The attention scores are then computed using hamilton product of G_q and C_k and then the score is applied to the quaternion values C_v:

$$Q_{Atten}^G = Q_{SM} \left(\frac{G_q \otimes C_k^*}{\|G_q\| \cdot \|C_k\|} \right) \otimes C_v \tag{3.7}$$

In this equation, $G_q \otimes C_k^*$ represents hamilton product of G_q with conjugate C_k^*, which can be used to capturing rotational alignment. $\|G_q\|$ and $\|C_k\|$ is Norms of G_q and C_k which provides normalization stability for quaternion interactions. Moreover, Q_{SM} is Quaternion-Softmax stabilizes the attention map, which is then applied to C_v to adjust the output feature maps accordingly. The significance of the quaternion attention map similar to real-valued attention mechanism. By incorporating the quaternion attention mechanism into the GB, the model enhances gradient features while ensuring that the key, value from color features of the CB helps in color awareness, improving the overall image enhancement in low-light conditions.

Building upon the attention mechanism within the CGE and CGD, the subsequent processing focuses on refining features. The gradient feature maps refined by the attention mechanism are processed through a quaternion convolution layer that doubles the number of feature maps followed by a depth-wise quaternion convolution layer. This dual-layered approach ensures robust extraction of structural details critical low-light conditions. Following this, the feature maps are chunked into illumination and reflectance components. The illumination component emphasizes intensity variations, while the reflectance component captures detailed textual features.

$$F_G^1, F_G^2 = \mathbb{Q}_{DQC}[\mathbb{Q}_{conv}(Q_{Attn}^G)] \tag{3.8}$$

A hyperbolic tangent activation maps the quaternion values to the range [-1, 1], enhancing contrast and capturing intricate intensity relationships, after the DQC on both F_G^1 and F_G^2 separately along with skip connections respectively. The recombination of illumination and reflectance via the Hamilton product results in a unified representation.

$$F_G^1 = tanh(F_G^1 + \mathbb{Q}_{DQC}(F_G^1))$$
$$F_G^2 = tanh(F_G^2 + \mathbb{Q}_{DQC}(F_G^2)) \tag{3.9}$$

$$I_g^{x+1} = Q_{Attn}^G + \mathbb{Q}_{conv}(F_G^1 \otimes F_G^2) \tag{3.10}$$

where I_g^{x+1} is the output feature map from the block. To further enhance edge sharpness, a skip connection is incorporated. This connection adds the original gradient feature maps directly to the refined output in an additive manner. The module preserves critical details from the initial gradient maps while focusing on edge regions. This ensures sharpness, more defined boundaries without sacrificing the integrity of the original features.

3.3 Gradient-aware Color Encoder (GCE) and Decoder (GCD) Block

The primary challenge in color enhancement under low-light conditions is the loss of vibrant and accurate color representation, which is exacerbated by low contrast and noisy

features. The Color Branch addresses this issue by incorporating the gradient information from the Gradient Branch, allowed the model to focus on enhancing color effectively in conjunction with edge details. In the CB, the GCE and GCD blocks processes the color features I_c^{inp}, estimated from the input image. The CB also benefits from feature sharing with the Gradient Branch. The quaternion query (C_q), is computed from the color features, I_c^x, while the quaternion key G_k and value G_v are derived from the gradient feature map, I_g^x. The sharing of features ensures that the color enhancement process is informed by the gradient related information from the GB, enabling better preservation of color and boundaries transitions.

The attention mechanism in the Color Branch operates also leverages cross quaternion attention with the color features used for estimating the quaternion query (C_q), while the gradient features are used for computing key (G_k) and value (G_v). The expression for quaternion cross attention is as follows:

$$C_q = \mathbb{Q}_{DQC}[\mathbb{Q}_{conv}(I_c^x)] \tag{3.11}$$

$$G_k, G_v = split[\mathbb{Q}_{DQC}(\mathbb{Q}_{conv}(I_g^x))] \tag{3.12}$$

Similar to GB, the quaternion convolution layer and depth-wise quaternion convolution layer ensure that spatial and channel-wise dependencies are captured efficiently, enabling effective attention to both color and gradient information. The attention score between C_q and G_k is calculated using the Hamilton product, and the final attention map is applied to the G_v as follows:

$$Q_{Atten}^C = Q_{SM}\left(\frac{C_q \otimes G_k^*}{\|C_q\| \cdot \|G_k\|}\right) \otimes G_v \tag{3.13}$$

This operation helps refine the color features in the CB while ensuring the gradient-based edge information influences the refinement. The feature sharing between branches ensures that the color features are enhanced not only based on their own characteristics but also with awareness of edge structures from the GB, leading to low-light enhancement task with better performance.

Following the attention mechanism, the refined features are further processed to ensure optimal enhancement of color components, particularly in challenging low-light scenarios. These features are passed through a quaternion convolution layer which doubles the feature map, followed by a depth-wise quaternion convolution layer, which extracts nuanced structural and intensity information while preserving multi-dimensional dependencies.

$$F_C^1, F_C^2 = \mathbb{Q}_{DQC}[\mathbb{Q}_{conv}(Q_{Attn}^C)] \tag{3.14}$$

Upon further enhancement of the representational fidelity, the processed feature maps are split into illumination and reflectance components. This splitting enables independent handling of light intensity and color details. A hyperbolic tangent activation function maps these quaternion values to the range [-1, 1], capturing complex intensity relationships and improving contrast. The illumination and reflectance are then recombined using the Hamilton product to produce an enhanced feature representation with refined color fidelity.

$$F_C^1 = tanh(F_C^1 + \mathbb{Q}_{DQC}(F_C^1))$$
$$F_C^2 = tanh(F_C^2 + \mathbb{Q}_{DQC}(F_C^2)) \tag{3.15}$$

$$I_c^{x+1} = \mathbb{Q}_{conv}(F_C^1 \otimes F_C^2) \tag{3.16}$$

Here I_c^{x+1} is the final output image from the block. The impact of this additional segment into the CGE-CGD and GCE-GCD is further explored and results are discussed in ablation study.

3.4   QLight-Net

Two primary issues often arise in enhancement task of low-light: first, the loss of edge information; and second, the inability to recover the true color of the objects. To address these challenges, the proposed architecture is designed with two parallel branches: the Gradient Branch (GB) and the Color Branch(CB). Each branch focuses on different aspects of the image. In the Gradient Branch, we apply a log-based gradient on the input image. This initial step is crucial for recovering subtle edges and boundaries in low-light conditions. The gradient information ensures that objects and regions within the image are properly delineated and can be process with better clarity. Subsequently, the Color Branch focuses on processing the color features of the image. By focusing on the key color components, the CB module ensures that the colors of regions in the image are preserved, leading to more realistic enhancement. Both branches work in parallel, with their outputs combined later in the network to generate a comprehensive enhancement of the original low-light image. This approach allows the targeted improvements in both edge clarity and color enhancement, addressing the core issues faced by low-light images in a balanced and efficient manner.

In the log-based gradient, gaussian smoothing reduces noise and softens edges in low-light images. The laplacian operator preserves edges by enhancing intensity shifts, preserving critical structural details for improved visibility. After extracting the log-based gradient feature map, which has a shape of $1 \times H \times W$, it directed to the GB after applying a multiple stacking operations. First, the gradient feature map is stacked three times and then the stacked feature maps is concatenated with the intensity channel I (C_3) from the channels generated by the Color Space Transform. This results in a 4-channel feature map that represents a single quaternion layer with the shape $4 \times H \times W$. This feature maps is then forwarded as input to the Gradient Branch. In parallel, the output from the Color Space Transform, which consists of a HVI color space image with dimensions $3 \times H \times W$, is concatenated with the log-based gradient feature map, resulting in a 4-channel feature maps representing color quaternion inputs of shape $4 \times H \times W$ to the Color Branch.

The GB and CB modules serve as the backbone of the proposed architecture, inspired by the U-Net framework. The simultaneous enhancement of both edge and color is crucial for accurate edges in low-light conditions. For effective feature enhancement, we utilize feature sharing between the two branches at each stage. This feature sharing is illustrated in 3.1, which depicts the connections between the GB and CB. The purpose of incorporating attention mechanisms in these modules is to address specific challenges in low-light images, where certain regions may be excessively dark, making it difficult to discern edges and accurately capture color information. To mitigate this, we apply targeted attention to these regions, gradually enhancing them at each step. This allows the model to focus on critical regions that require more attention, progressively improving the enhancement results. To achieve this, the Gradient Branch and Color Branch are consists of six stages, where each stage builds upon the previous one, progressively enhancing the image.

In both branches, it applies a single layer of quaternion convolution on the respective inputs to the GB and CB, resulting in feature maps of size $C \times H \times W$ to learn the initial features and also to map the features to quaternion representations. In the first

stage of encoder modules, it performs a down-sample operation on the feature maps and transforms it to $C \times \frac{H}{2} \times \frac{W}{2}$, and directed to the CGE1 and GCE1 in the respective branches after transmitting a residual connection-1. Before the second stage, it performs down-sampling operation after transmitting a residual connection-2. The input to the CGE2 and GCE2 modules is transformed to $2C \times \frac{H}{4} \times \frac{W}{4}$. And the input to CGE3 and GCE3, after residual connection-3 and down-sampling operation that alters the shape to $4C \times \frac{H}{8} \times \frac{W}{8}$. The output from the CGE3 and GCE3, the shape is retained and connections are forwarded to the decoder-half of the network. From CGD3 and GCD3, image restoration is initiated after the corresponding decoder modules are activated using the residual connections. These residual connections allow progressive refinement of feature maps, ensuring that important details are retained while enhancing the image structure. The output shape of CGD3 and GCD3 is $4C \times \frac{H}{8} \times \frac{W}{8}$, which will be transformed to $2C \times \frac{H}{4} \times \frac{W}{4}$ via up-sampling and concatenation with residual connection-3. Similarly, CGD2-GCD2 and CGD1-GCD1 will perform the same set of corresponding operations and the final output will be retransformed to the shape $C \times H \times W$. At the final stage, the enhanced feature maps from both the GB and CB, will be directed to the last single layer of quaternion convolution which will revert back the quaternion representations to the real-valued number. The final output from both the branches will be a 1-channel and 2-channel feature maps which represent the modified color space and are concatenated to produce a 3-channel final HVI output from the branch. Additionally, a residual connection is established by adding the HSV image output from the Color Space Transform, ensuring that the integrity of the image's details and structure is preserved.

3.5   Loss Function

A loss function is equally important in an architecture, which helps the modules lo learn efficiently and effectively. In the proposed architecture we inducted three main loss components. First L1 loss to maintain the average absolute difference between the enhanced image and ground truth. Since the architecture incorporates gradient feature map and edge is one of the main factor of this proposed method so the losses on edges to preserves the structural loss, along with the perceptual loss for high level features of the enhanced image.

Since, the input image is in RGB color space and then converted to modified and updated HSV color space, so the error rate or loss on conversion and handling needs to be taken care off. Gradient Feature Map is also used to as a input to the GB, losses on edge needs to be considered along with perceptual loss for both HVI color space image as well as RGB color space image.

$$L = L_1(\hat{I}_{en}, I_{GT}) + L_{edge}(\hat{I}_{en}, I_{GT}) + L_p(\hat{I}_{en}, I_{GT}) \qquad (3.17)$$

Here, $\hat{I}_{en}$ and I_{GT} represents the enhanced image and ground truth image respectively. And, L_1 is L1 loss, L_{edge} is about the edge loss and L_p is perceptual loss.

We use the pixel-wise L1 loss to minimize the average absolute difference between the enhanced output $\hat{I}en$ and the ground truth image I_{GT}. This ensures overall luminance and color fidelity. Given that edge preservation is crucial especially when using gradient-based features and operating in the modified HSV color space, we include an edge loss

to maintain structural consistency. The Sobel operator is applied to extract edge maps from both $\hat{I}en$ and I_{GT}, and the L1 loss between them is computed. To capture high-level semantic features and ensure perceptual similarity, we utilize a perceptual loss based on a pretrained VGG network.

Chapter 4

RESULTS and DISCUSSION

4.1 Implementation Details

The following experiments are conducted on Ubuntu 22.04 equipped system with a GPU of 24 GB. With the batch size set to 16, the model was trained to 7500 epochs on the LOLv1 and LOLv2 datasets. Using the patch size 256×256 cropped 10 samples from each training image to ensure robust generalization. LOLv2 dataset is further is of two type: Synthetic and Real. The training, validation and testing of our method follows the similar splits as given in the official dataset. The results of these experiments are presented in detail in subsequent sections on ablation studies, quantitative and qualitative sections below.

4.2 Ablation Study

Extensive experiments are being conducted to validate the proposed architecture, which consists of several critical components. We designed two focused ablation studies to understand the contribution of each module. As part of the first study, the impact of proposed Depth-wise Quaternion Convolution in the architecture and it's role in learning. In the second ablation, it evaluates the impact of our dual-branch architecture, analyzing how the Gradient Branch (GB) and the Color Branch (CB) work independently in the absence of other branch. This section provides a detailed analysis of these studies, along with corresponding results to demonstrate the significance of these modules.

4.2.1 Ablation 1 - Depth-wise Quaternion Convolution

In this section, we present an ablation study to assess the impact of the proposed Quaternion Depth-wise Convolution (DQC) layer in terms of both performance and parameter efficiency. The main advantage of Depth-wise quaternion convolution is its capacity to apply separate filters to every feature map. This preserves the model's capacity to efficiently process high-dimensional input while drastically lowering the number of parameters. The ablation analyses the performance of using DQC in terms of computing efficiency and performance metrics. To evaluate the impact of the DQC, we conducted experiments under three configurations: first, without DQC layer to assess the model's performance; second, by replacing DQC with a quaternion convolution layer to understand the effects of using traditional quaternion convolution; and finally, by incorporating the quaternion depth-wise convolution layer into our proposed model to analyze its overall contribution to performance.

Configurations	PSNR ↑	SSIM ↑	LPIPS ↓
w/o Depth-wise Quaternion Convolution	26.57	0.858	0.085
w/ Quaternion Convolution	28.09	0.875	0.056
w/ Depth-wise Quaternion Convolution	29.05	0.880	0.047

Table 4.1: Ablation study results for the impact of the Depth-wise Quaternion Convolution layer. Comparison of PSNR, SSIM, and LPIPS across various configurations

As shown in 4.1, it is clearly evident that removing the depth-wise convolution results in a considerable drop in all performance metrics, indicating that this layer plays a crucial role in learning discriminative features in low-light conditions. Although the use of traditional quaternion convolution improves performance compared to the baseline, it comes at a significant computational cost. The proposed DQC layer, however, not only yields superior results in evaluation of PSNR, SSIM, and LPIPS but is also parameter efficient.

DQC	Quaternion Conv	Conv	PSNR ↑	SSIM ↑	LPIPS ↓	Parameters(M)
✗	✗	✓	28.14	0.889	0.079	100.83
✗	✓	✗	28.09	0.875	0.056	25.314
✓	✓	✗	29.05	0.880	0.047	0.747

Table 4.2: Ablation study results for number of parameters with DQC and without DQC. Comparison of parameters (in millions) across various configurations.

To further analyze the parameter reduction claim, we perform a comparison of the number of trainable parameters under various combinations of convolutional modules, as shown in 4.2. The configuration with only standard convolution achieves decent performance but suffers from a very high parameter count (100.83 million). On the other hand, replacing it with traditional quaternion convolution reduces parameters significantly to 25.314 million, while maintaining acceptable performance. However, our proposed configuration that uses the Quaternion Depth-wise Convolution layer not only outperforms both configurations in terms of image quality but also drastically reduces the parameter count to just 0.747 million.

It is evident from the ablation that the DQC layer provides a trade-off between performance and model complexity. By integrating this module into the network, we achieve state-of-the-art enhancement results while maintaining a lightweight architecture.

4.2.2 Ablation 2 - Impact from Dual Branch

Our architecture incorporates a dual-branch design: Gradient Branch and Color Branch. Both branches share information, ensuring mutual learning and contributing to robust image enhancement. In this study, we evaluate the impact of these branches individually and in combination on the final results as shown in 4.3.

We perform three experiments to analyze the roles of the two branches: first, retaining only Color Branch to observe its ability to compensate for edge enhancement in the absence of Gradient Branch; second, retaining only Gradient Branch to evaluate its impact on color correction and refinement; and third, using both Gradient and Color Branches together to assess their combined effectiveness.

The results demonstrate the significance of both branches in the proposed architecture. The exclusion of Gradient Branch results in a PSNR of 27.92, SSIM of 0.860, and LPIPS

GB	CB	PSNR ↑	SSIM ↑	LPIPS ↓
✗	✓	27.92	0.86	0.075
✓	✗	28.10	0.861	0.078
✓	✓	29.05	0.880	0.047

Table 4.3: Ablation study results for the impact of the Gradient Branch (GB) and Color Branch (CB) on image quality. Comparison of PSNR, SSIM, and LPIPS with and without the individual branches and both branches combined

of 0.075, indicating a decline in edge-focused learning, highlighting GB's role in edge enhancement. Similarly, excluding Color Branch yields PSNR of 28.10, SSIM of 0.861, and LPIPS of 0.078, reflecting its critical contribution to color correction and the preservation of natural colors. When both Gradient Branch and Color Branch are included, the architecture achieves the best performance with PSNR of 29.05, SSIM of 0.880, and LPIPS of 0.047, confirming that the two branches collaboratively enhance the overall image quality while addressing distinct aspects of edge and color refinement.

4.2.3 Ablation 3 - Impact of different input to GB of our method

To evaluate the impact of different input channel configurations on the performance of our model, we conduct an ablation study focusing on the assignment of the real and imaginary components (r, $\hat{i}$, $\hat{j}$, $\hat{k}$) in the quaternion representation, specifically in the Gradient Branch (GB) of our proposed QLight-Net. We compare three different configurations of quaternion inputs and report their performance in terms of PSNR, SSIM, and LPIPS metrics. The results are summarized in Table 4.4. The performance of our model is also included for comparison.

Configurations	PSNR ↑	SSIM ↑	LPIPS ↓
Configuration 1	23.74	0.850	0.126
Configuration 2	25.67	0.855	0.066
Configuration 3	28.52	0.873	0.051
Our Method	29.05	0.880	0.047

Table 4.4: Ablation study of various configurations with different input to the gradient branch of the proposed model.

In Configuration 1, the real part r is assigned the RGB gradient of the input image, while the imaginary parts $\hat{i}$, $\hat{j}$, and $\hat{k}$ correspond to the gradients of the intensity channel along the x-axis, y-axis, and combined x-y axes, respectively. This configuration provides suboptimal performance. Configuration 2 assigns all four components (r, $\hat{i}$, $\hat{j}$, and $\hat{k}$) with gradient maps derived from the intensity channel of the HVI color space. Although it performs better than Configuration 1, the absence of color gradient information leads to limited structural recovery. In Configuration 3, the real and first imaginary parts are assigned to the intensity channel of the HVI image, while the remaining components are set as RGB gradients. This configuration results in a noticeable improvement. Finally, in our proposed method, we assign the gradient of the RGB image uniformly across all quaternion components. The proposed method achieves the best performance across all evaluation metrics (PSNR, SSIM, and LPIPS), indicating that the RGB gradient

information enriches the spatial detail and structural integrity of the enhanced image. Moreover, we observed that the quaternion input configuration in the Color Branch had minimal effect on the enhanced results, as the network is capable of learning inter-channel correlations during training.

4.2.4 Ablation 4 - Effect of Diverse Information Sources in Quaternion Representation and Input to GB-CB

As discussed in 3, our proposed architecture consists of two branches: the Gradient Branch and the Color Branch. Both the branches are responsible for different tasks while sharing information. Each of these branches processes 4-channel inputs, which are then forwarded into a single layer of quaternion convolution layer. These 4-channels are treated as the $\mathbf{r}$, $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$ components of a quaternion and are further processed accordingly. In this section, we analyze the impact of different channel combinations for the quaternion representation. The channels in the arrangement to the $\mathbf{r}$, $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, $\hat{\mathbf{k}}$ components can influence the learning process. With the studies and experimenting, we aim to understand the set of channels that yield better performance.

Configurations	Gradient Branch	Color Branch	PSNR↑	SSIM↑	LPIPS↓
Configuration 1	**C1:** Gradient RGB **C2:** Gradient along x-axis of Intensity channel **C3:** Gradient along y-axis of Intensity channel **C4:** Gradient along x- and y-axis of Intensity channel	**C1C2C3:** HVI **C4:** Gradient RGB	23.74	0.850	0.126
Configuration 2	**C1:** Gradient along I-channel of HVI **C2:** Gradient along I-channel of HVI **C3:** Gradient along I-channel of HVI **C4:** Gradient along I-channel of HVI	**C1C2C3:** HVI **C4:** Gradient RGB	25.67	0.855	0.066
Configuration 3	**C1:** I-channel of HVI **C2:** I-channel of HVI **C3:** Gradient RGB **C4:** Gradient RGB	**C1C2C3:** HVI **C4:** Gradient RGB	28.52	0.873	0.051

Table 4.5: Performance analysis when each channel in the Gradient Branch and Color Branch contains heterogeneous information, combining different gradient and color components in a single quaternion representation.

In 4.5 it presents a set of channels for the Gradient Branch and Color Branch in our proposed architecture. In the Gradient Branch, the set of four channels: the first channel is the log-based gradient applied on input RGB image, while the second, third and fourth channel are derived from the I-channel of HVI color space, where a log-based gradient is applied along the x-axis, y-axis and both x- and y-axes, respectively. Similarly, for the Color Branch it takes the HVI channels along with a log-based gradient computed directly from the input RGB image as the fourth channel. Gradient Branch is designed to focus on edge information, it leverages the log-based gradient features extracted at different levels to enhance feature representation within the quaternion framework. While the Color Branch benefits from color information and concatenation with gradient features. However, the results suggest from 4.5 that the model struggles with this set of channels combination. The model may struggle to effectively fuse such diverse directional information into a single quaternion representation. Quaternion-based processing enhances efficiency, but representation of real and imaginary values in certain channel configurations may pose challenges in achieving an optimal feature representation.

4.2.5 Ablation 5 - Impact of Homogeneous Information in Quaternion Components and Input to GB-CB

With an update in architecture, 4.6 analyzes the impact of modification on the input to Gradient Branch.

Gradient Branch	Color Branch	PSNR↑	SSIM↑	LPIPS↓
C1: Gradient along I-channel of HVI **C2:** Gradient along I-channel of HVI **C3:** Gradient along I-channel of HVI **C4:** Gradient along I-channel of HVI	**C1C2C3:** HVI **C4:** Gradient RGB	25.67	0.855	0.066

Table 4.6: Impact of using homogeneous information within each quaternion representation by assigning identical log-based gradient features across all channels in the Gradient Branch, while keeping the Color Branch unchanged.

In the 4.5, it can be observed a significant drop in performance when incorporating diverse types of channels within a single quaternion representation. This indicates heterogeneous information sources in quaternion components could negatively affect the learning process. To address this, we configured a strategy where all four channels in the Gradient Branch. Specifically, we applied a log-based gradient to the I-channel of the HVI color space, and concatenated 4-times in a stack. Meanwhile, the Color Branch remains the same configuration as in 4.5. This modification led to a noticeable increase in PSNR and a decrease in LPIPS, indicating that providing similar information across quaternion channels enables the model to learn more effectively.

4.2.6 Ablation 6 - Balancing Structural and Gradient Information for Optimal Representation and Input to GB-CB

From both the prior methods, it is observed that channels with similar base information within a single quaternion representation is more effective for learning. Building upon this understanding, we configured the experiment with the Gradient Branch with two distinct types of information, however within each group both the channels contain homogeneous information. The first is the intensity of each pixel, which may contribute to detecting edges by capturing intensity variations. The second is the log-based Gradient applied to the RGB image, which enhances edge information extraction.

Gradient Branch	Color Branch	PSNR↑	SSIM↑	LPIPS↓
C1: I-channel of HVI **C2:** I-channel of HVI **C3:** Gradient RGB **C4:** Gradient RGB	**C1C2C3:** HVI **C4:** Gradient RGB	28.52	0.873	0.051

Table 4.7: Effect of mixed homogeneous groups, where each quaternion in the Gradient Branch consists of two distinct but internally homogeneous types of information: intensity-based features and log-based gradient applied to the RGB image.

As quantitative results shown in 4.7, this approach leads to an improvement in structural quality, with a significant impact attributed to the log-based gradient on the RGB

image. We can conclude the observation that incorporating log-based gradient information from the RGB domain plays a crucial role in refining the final structure, making it a highly influential component in our architecture.

4.3   Computational Efficiency Comparison

Model efficiency is an important aspect for real-world adaptability. Achieving superior enhancement quality is essential, but a model with higher parameter count becomes computationally expensive. In this proposed architecture, we address this issue by leveraging a quaternion-based architecture, which maintains a significantly lower computational efficiency as well as state-of-the-art quantitative results. Our focus is to achieve competitive or superior performance while requiring substantially lower parameter count, demonstrating the efficiency of the proposed approach.

The quaternion representation plays a pivotal role in this proposed method as discussed in the methodology section. The inherent advantage allows our model to process information more effectively, reducing redundancy and computational overhead. To quantify the efficiency of our model, we compare it with existing methods on three primary performance metrics as discussed above in 4.4. Many models achieve comparative high performances but come with parameter counts in millions, making them computationally expensive. In contrast, our model achieves comparable or superior results with 0.747 million parameters, making it highly efficient even on standard GPU hardware configuration.

For a comprehensive comparison, 4.8 provides a detailed analysis. Additionally 4.1 illustrates the efficiency comparison across all three datasets. Specifically, we plot three key efficiency graphs:

- PSNR vs. Parameters (in millions)4.1a: The x-axis represents the parameter count, while the y-axis denotes the PSNR value. A model positioned closer to top-left signifies superior efficiency, achieving a high PSNR with minimal parameters.

- SSIM vs. Parameters (in millions)4.1b: The x-axis represents the parameter count, while the y-axis indicates the SSIM value. Similar to PSNR, a model closer to top-left demonstrates the best trade-off between performance and efficiency. As high the SSIM value to achieve with lesser number of parameter.

- LPIPS vs. Parameter (in million)4.1c: The x-axis represents the parameter count, while the y-axis denotes the LPIPS score. In this case, a model closer to the bottom-left is preferred, as it achieves a lower LPIPS value with fewer parameters.

From these graphical comparisons, it is evident that our proposed model consistently outperforms existing approaches in terms of efficiency while maintaining state-of-the-art enhancement quality.

4.4   Quantitative Evaluation

To compare the proposed architecture with state-of-the-art methods, it uses 3 primary performance metrics: PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index Measure) and LPIPS (Learned Perceptual Image Patch Similarity). These metrics assess the performance of the architecture from different perspectives, helping to evaluate its effectiveness and enabling comparisons with existing methods.
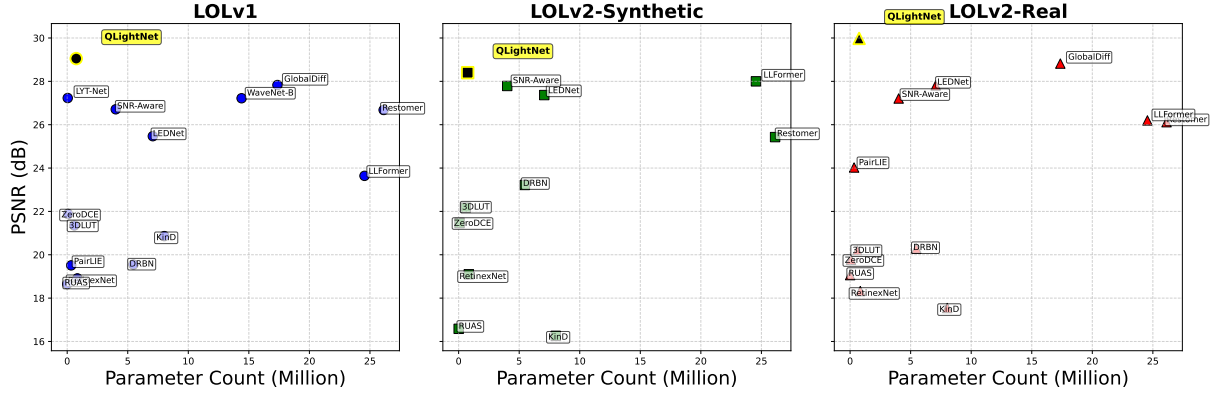
(a) PSNR(dB) vs. Parameter Count(in millions) for various models for different datasets.
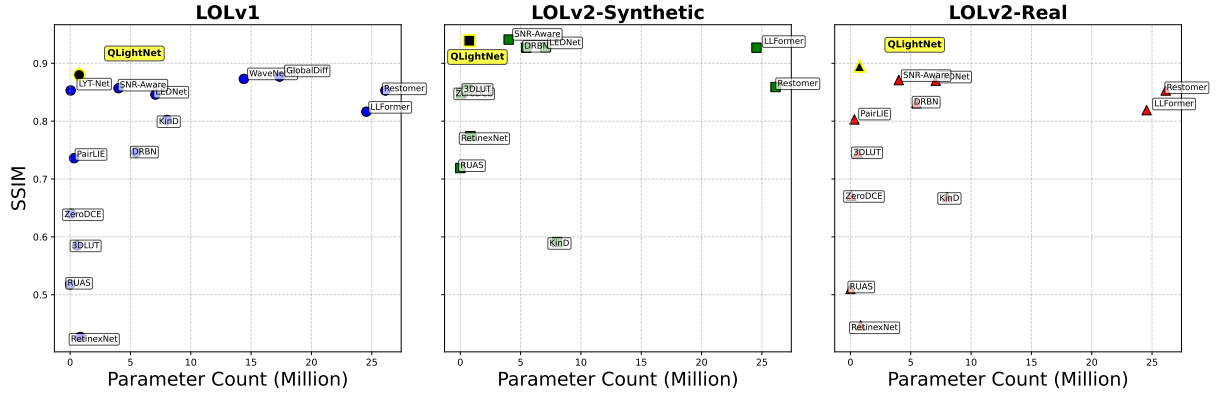
(b) SSIM vs. Parameter Count(in millions) for various models for different datasets.
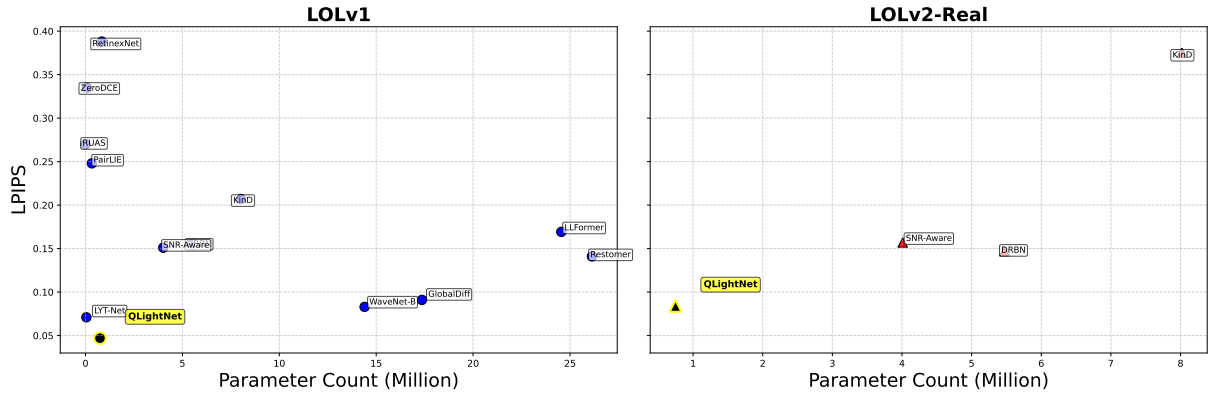
(c) SSIM vs. Parameter Count(in millions) for various models for different datasets.

Figure 4.1: Comparison of (a)PSNR(dB) vs. Parameter Count(in millions) (b)SSIM vs. Parameter Count, (c)LPIPS vs. Parameter Count for various low-light image enhancement models across LOLv1(left), LOLv2-Synthetic(middle) and LOLv2-Real(right) datasets. The proposed method QLightNet shown with a highlighted yellow tag.

Methods	LOLv-1			LOLv-2(Synthetic)			LOLv-2(Real)			Params/M
	PSNR(dB)↑	SSIM↑	LPIPS↓	PSNR(dB)↑	SSIM↑	LPIPS↓	PSNR(dB)↑	SSIM↑	LPIPS↓	
RetinexNet[1] (CVPR' 18)	18.91	0.427	0.388	19.09	0.774	-	18.32	0.447	-	0.84
KinD[47] (CVPR' 19)	20.86	0.802	0.207	16.25	0.591	-	17.54	0.669	0.375	8.02
ZeroDCE[13] (CVPR' 20)	21.88	0.640	0.335	21.46	0.848	-	19.77	0.671	-	0.075
3DLUT[48] (IVP' 20)	21.35	0.585	-	22.17	0.854	-	20.19	0.745	-	0.59
DRBN[24] (CVPR' 20)	19.55	0.746	0.155	23.22	0.927	-	20.29	0.831	<u>0.147</u>	5.47
EnlightenGAN[25] (TIP' 21)	20.00	0.691	0.322	16.57	0.734	-	18.23	0.617	-	114.35
LEDNet[26] (ECCV' 22)	25.47	0.846	-	27.36	0.928	-	27.81	0.870	-	7.07
KinD++[4] (IJCV' 21)	19.51	0.736	0.248	-	-	-	24.02	0.803	-	8.27
LLFormer[29] (AAAI'23)	23.64	0.816	0.1692	28.00	0.927	-	26.19	0.819	-	24.55
RetinexFormer[5] (CVPR' 23)	25.16	0.845	0.850	**28.99**	**0.939**	-	27.69	0.856	-	1.61
WaveNet-B[6] (PG' 23)	27.22	0.873	<u>0.083</u>	-	-	-	-	-	-	17.42
EMNet[3] (TMM' 23)	25.37	0.868	0.086	-	-	-	-	-	-	12.52
GlobalDiff[15] (CVPR' 23)	<u>27.83</u>	<u>0.877</u>	0.091	-	-	-	<u>28.82</u>	-	-	17.36
RLLIE[38] (TIM' 24)	20.12	0.770	0.220	-	-	-	-	-	-	-
YUVAtten.Net[37] (CG' 25)	22.74	0.846	0.079	21.11	<u>0.931</u>	-	20.48	0.849	-	0.12
BiFormer-B[7] (TMM' 25)	24.03	0.856	0.110	24.81	0.928	<u>0.065</u>	22.93	0.860	0.149	1.43
Ghillie-MD[36] (TCSVT' 25)	23.87	0.835	0.132	-	-	-	-	-	-	3.05
Our Method	**29.05**	**0.880**	**0.047**	<u>28.40</u>	**0.939**	**0.050**	**29.98**	**0.895**	**0.084**	0.747

Table 4.8: Comparison of Proposed method With the State-of-the-Art Low Light enhancement Methods on LOL-v1[1], LOL-v2[2] with two versions of Synthetic and Real ordered in timeline of publication. Comparing PSNR(↑), SSIM(↑) and LPIPS(↓). ↑ represents Higher the better and ↓ represents Lower the better. **Bold** represents the best value and <u>UNDERLINE</u> represents second-best result.

We present a comparative analysis of the performance of proposed model against several state-of-the-art methods as shown in 4.8. Higher values of PSNR and SSIM indicate better performance, whereas lower LPIPS values are preferable for improved perceptual quality. We conduct experiments on three benchmark datasets: LOLv1, LOLv2-Synthetic, and LOLv2-Real. The results for all competing methods where other models are compared, and our proposed model is evaluated alongside these methods.

On the LOLv1 dataset, our model achieves the best performance among all evaluated methods, outperforming the other models in terms of all the considered metrices. Specifically, the proposed model attains a SSIM score of 0.880, PSNR of 29.05dB, and LPIPS score of 0.047, which represents the highest values among all the compared models. In terms of quantitative results, our model outperforms many models which are based on various techniques with fewer parameter. It is noteworthy that our model is computationally efficient, achieving state-of-the-art results with only 0.747 million parameters.

For the LOLv2-Synthetic dataset, the proposed model continues to exhibit competitive performance, achieving the best scores on SSIM of 0.939 and LPIPS score of 0.050. While the PSNR of 28.40dB places our model second best. Despite having a relatively small number of parameters, the model still demonstrates the ability to achieve impressive results on synthetic low-light data. This indicates that the proposed method is effective across various types of low-light image data, even when dealing with synthetic datasets.

On the more challenging LOLv2-Real dataset, which consists of real-world low-light images, our method significantly outperforms competing approaches. It achieves the best scores in all three metrics which check different aspect of an enhanced image. With PSNR value of 29.98dB and the SSIM score of 0.895 being the highest among all the comparing methods with lowest LPIPS score of 0.084. These results highlighting the robustness of our method in preserving fine image details and maintaining structural integrity in real-world low-light conditions.

In summary, our proposed model demonstrates superior performance across all three datasets when evaluated using objective metrics (PSNR, SSIM, and LPIPS). Despite the relatively small number of parameters (0.74M), our approach offers state-of-the-art performance with significantly reduced computational complexity. This validates the effectiveness and efficiency of our method for low-light image enhancement, making it a promising solution for both synthetic and real-world low-light scenarios.

The results of the proposed architecture, compared to other models, are detailed in 4.8. It is evident that our model outperforms many state-of-the-art models in multiple metrics. From another perspective, considering the significantly lower parameter count, the proposed model achieves results comparable to the leading methods in the field. Specifically, when trained on the LOLv1 dataset, our model performs better than the best model in PSNR, SSIM and LPIPS metrics and ranks the best models. On the LOLv2-Synthetic dataset, our model achieves best rankings in both SSIM and LPIPS, and provides the second-best performance in PSNR. In the case of LOLv2-Real, the proposed model surpasses the best model in PSNR, SSIM, and LPIPS. When considering the parameter count, it is clear that the proposed architecture achieves excellent results with significantly fewer parameters compared to other models. This trade-off between performance and parameter efficiency makes our approach highly competitive, offering near-best results while being computationally efficient.

4.5 Qualitative Evaluation

In addition to quantitative evaluation, visual assessment plays a crucial role in evaluating the perceptual quality of low-light image enhancement.

The proposed architecture enhances both image quality and structural consistency effectively. In 4.2, we show a visual comparison of the output images generated by our model with the input and the ground-truth. It is evident from these results that our model is capable of producing high-quality images that closely resemble the ground truth, demonstrating its effectiveness in enhancing low-light images. The enhanced images retain key details, sharpness, and clarity, which are critical for perceptual improvement in low-light scenarios. The primary goal of our model is not only to enhance the image but also to preserve the original color information and structural integrity. The comparison clearly highlights our model's superior performance in this regard. Upon zooming into the enhanced regions, it becomes apparent that our method avoids color distortion and over-enhancement, preserving natural color tones and accurately matching the ground-truth
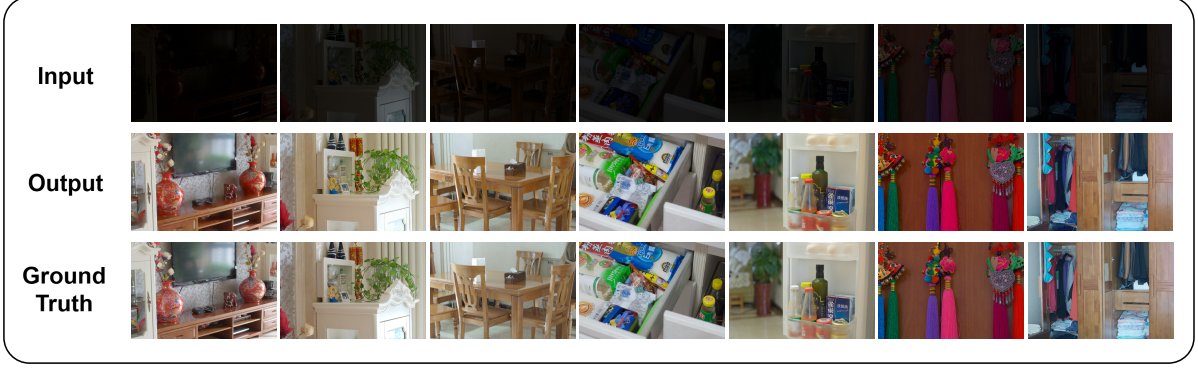
Figure 4.2: Visual output from the proposed architecture on various training images from LOLv1 dataset.
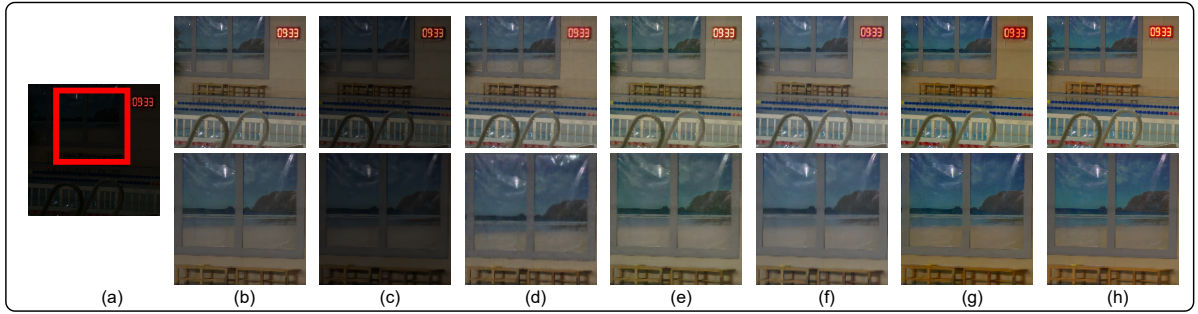
image.

Figure 4.3: Visual results with zoomed-in marked red section of various methods for "Pool" image. (a) input images; (b) EMNet[3]; (c) KinD++[4]; (d) RetinexFormer[5]; (e) WaveNet-B[6]; (f) BiFormer-B[7]; (g) Our method; and (h) Ground Truth.

Our proposed architecture not only enhances image quality but also maintains structural consistency, effectively improving the visual appeal of enhanced images. Visual results of various methods are given in 4.3 and 4.6 to compare the enhancement by our model with other contemporary methods. The visual result contains enhanced images in the top row in each figure and enlarged section of the enhanced result marked in the input image in the bottom row for better clarity.

The results of EMNet show lightness enhancement with limited contrast. In addition, it is unable to retain the appropriate colors as shown in 4.3(b), 4.4(b), 4.5(b), and 4.6(b). It can be observed from 4.3(c) to 4.5(c) that KinD++ is not able to properly enhance the lightness. Moreover, the lightness in the results shown in 4.4(c) and 4.6(c) is enhanced; however, the details and color are not significantly enhanced. The results of RetinexFormer enhance adequate lightness; however, the details in the enhanced images are not clear due to over smoothing as shown in 4.3(d), 4.4(d), 4.5(d) to 4.6(d). WaveNet enhances details and lightness as shown in 4.3(e), 4.5(e), and 4.6(e). However, WaveNet provides faded colors in the enhanced images. In 4.3(f), 4.4(f), 4.5(f), and 4.6(f), the results of BiFormer-B show faded colors and smooth details which results in poor enhancement quality. The enhanced images of QlightNet are given in 4.3(g), 4.4(g), 4.5(g), and 4.6(g). The results are quite closer to the ground-truth images in terms of details, contrast, and colors. The result shown in 4.5(g) shows slightly better colors than the
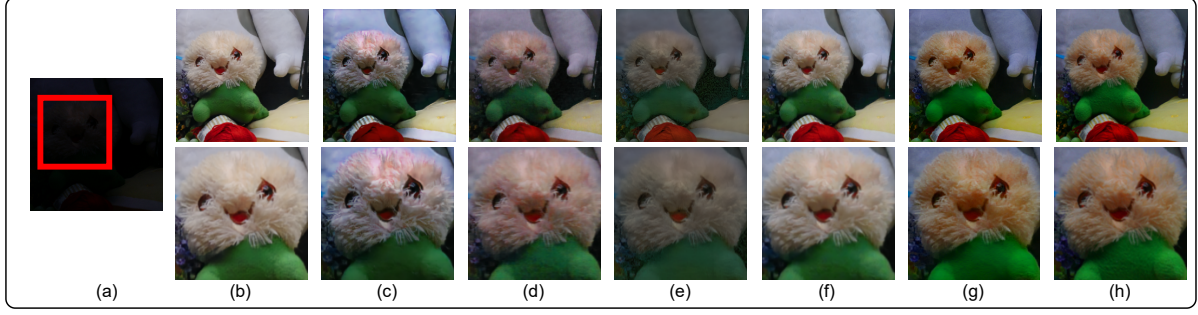
Figure 4.4: Visual results with zoomed-in marked red section of various methods for "Teddy" image. (a) input images; (b) EMNet[3]; (c) KinD++[4]; (d) RetinexFormer[5]; (e) WaveNet-B[6]; (f) BiFormer-B[7]; (g) Our method; and (h) Ground Truth.
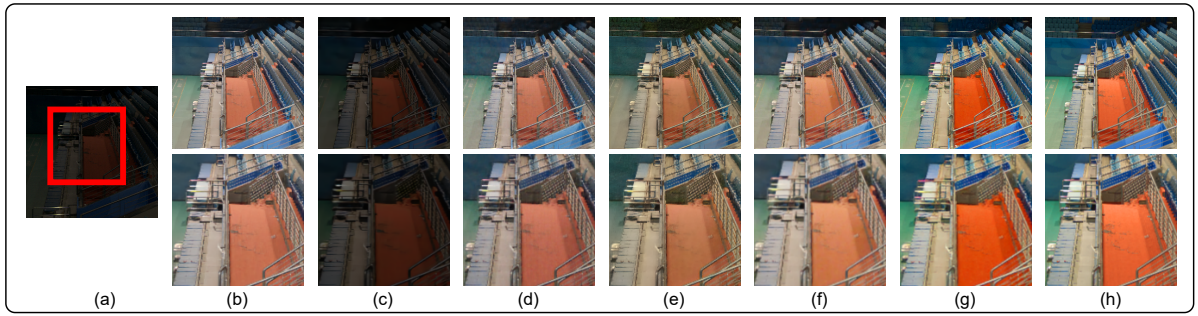
Figure 4.5: Visual results with zoomed-in marked red section of various methods for "Stands" image. (a) input images; (b) EMNet[3]; (c) KinD++[4]; (d) RetinexFormer[5]; (e) WaveNet-B[6]; (f) BiFormer-B[7]; (g) Our method; and (h) Ground Truth.
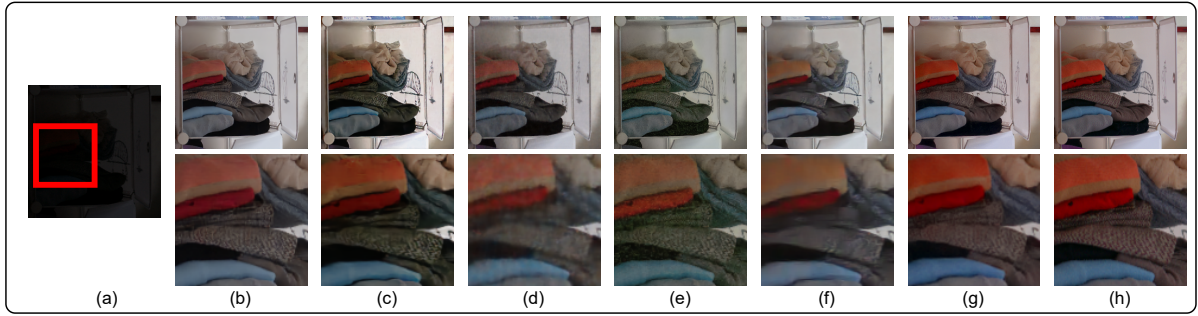
Figure 4.6: Visual results with zoomed-in marked red section of various methods for "Shelf" image. (a) input images; (b) EMNet[3]; (c) KinD++[4]; (d) RetinexFormer[5]; (e) WaveNet-B[6]; (f) BiFormer-B[7]; (g) Our method; and (h) Ground Truth.

ground truth image, which shows the details quite clearly in the image. These visual results demonstrate the effectiveness of the proposed method, showing that it not only enhances the visual quality of low-light images but also maintains the color and structures that are essential for improving perceptual quality in various applications.

4.6 Time Comparison

This section presents a comparative analysis of our proposed QLightNet with respect to parameters and inference time. Table 4.9 shows the number of parameters (in millions) and the average inference time (in seconds) for images of resolution 256×256, across several state-of-the-art methods.

Methods	Parameters(millions)	Time(sec)
KinD++[4]	8.27	0.13
RetinexFormer[5]	1.61	0.35
WaveNet-B[6]	17.42	0.12
EMNet[3]	12.52	0.13
YUVAtten. Net[37]	0.12	0.19
BiFormer-B[7]	1.43	2.08
Ghillie-MD[36]	3.05	0.42
Our Method	0.747	0.09

Table 4.9: Comparison of parameter count and inference time of images across various state-of-the-art methods.

It can be noted from table 4.9, our method achieves a parameter count of only 0.747 million, which is significantly lower than many contemporary methods. For instance, KinD++ [4] and EMNet [3] require 8.27M and 12.52M parameters respectively, whereas WaveNet-B [6] requires a heavy computational footprint with 17.42M parameters. Even lightweight models such as RetinexFormer [5] and BiFormer-B [7] remain above the 1 million parameter mark.

Moreover, our model also demonstrates favorable inference efficiency, achieving an average processing time of 0.09 seconds per image, which is significantly faster than most existing models. Few methods like WaveNet-B [6] and KinD++ [4] show comparable speed (0.12–0.13 seconds); however, they do so at the cost of significantly higher parameter counts. On the other hand, YUVAtten. Net [37] achieves lower parameter count (0.12M) but at the expense of increased inference time (0.19 seconds). Our method achieves a balanced and efficient design, leveraging quaternion convolution to compress feature representations without degrading inference speed. This shows the improved computational efficiency of our proposed method.

Chapter 5

CONCLUSION AND FUTURE SCOPE

This paper presents a new quaternion based light-weight network for low-light image enhancement. The proposed QLight-Net uses our novel depth-wise quaternion convolution that captures intra-channel spatial features and inter-channel dependencies. Moreover, we proposed a two-branch model that leverages the developed quaternion cross-attention for encoder and decoder. The proposed model extracts features using gradient and color branches to effectively enhance low-light images. The proposed method outperforms most contemporary methods in quantitative and visual analysis. Furthermore, the proposed method uses a lesser number of parameters to outperform other methods.

However, the proposed approach has certain limitations. The model is trained on static image datasets and may not generalize well to dynamic scenes or video data. Also, in cases of extremely low illumination or presence of severe artifacts such as motion blur, the enhancement quality may degrade. In future work, we intend to extend the model to video-based low-light enhancement by incorporating temporal information. We also aim to explore domain adaptation techniques for better generalization across diverse scene types. Furthermore, integrating self-supervised or semi-supervised learning methods could help reduce the dependency on large-scale paired datasets.

Our method achieves superior enhancement in low-light images, but certain limitations exist. A minor limitation of our method is the slight over-enhancement of certain color components, such as orange hues, in some particular cases. This could arises due to inherent nature of quaternion-based dedicated branch for color modeling. Future work can focus in refining adaptive color correction strategies within the quaternion space to mitigate this effect while preserving the advantages of quaternion-based feature learning.

In the main proposed architecture, the two branches parallel learning in different aspect in intensity feature map and color feature map, and in the end concatenate the information to form a final enhanced image. Since it is extracting the color feature map and intensity feature map separately, so it need to be combined again in the same structure as input image. Post processing restructure the enhanced image to reshaped it into the same structure as of input image. From the results it can be concluded that the proposed method has tried to provide a very optimal architecture to maintain the model lightweight as well as able to achieve good results as compare to many state-of-the-art methods.

Bibliography

[1] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in *British Machine Vision Conference*, 2018.

[2] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3291–3300.

[3] D. Ye, Z. Ni, W. Yang, H. Wang, S. Wang, and S. Kwong, "Glow in the dark: Low-light image enhancement with external memory," *IEEE Transactions on Multimedia*, pp. 1–16, 2023.

[4] Y. Zhang, X. Guo, J. Ma, W. Liu, and J. Zhang, "Beyond brightening low-light images," *International Journal of Computer Vision*, vol. 129, pp. 1013–1037, 2021.

[5] Y. Cai, H. Bian, J. Lin, H. Wang, R. Timofte, and Y. Zhang, "Retinexformer: One-stage retinex-based transformer for low-light image enhancement," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2023, pp. 12 504–12 513.

[6] J. Dang, Z. Li, Y. Zhong, and L. Wang, "WaveNet: Wave-Aware Image Enhancement," in *Pacific Graphics Short Papers and Posters*, R. Chaine, Z. Deng, and M. H. Kim, Eds. The Eurographics Association, 2023.

[7] R. Xu, Y. Li, Y. Niu, H. Xu, Y. Chen, and T. Zhao, "Bilateral interaction for local-global collaborative perception in low-light image enhancement," *IEEE Transactions on Multimedia*, 2024.

[8] C. Li, C. Guo, L. Han, J. Jiang, M.-M. Cheng, J. Gu, and C. C. Loy, "Low-light image and video enhancement using deep learning: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 12, pp. 9396–9416, 2022.

[9] R. C. Gonzales and P. Wintz, *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc., 1987.

[10] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Am.*, vol. 61, no. 1, pp. 1–11, Jan 1971. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=josa-61-1-1

[11] K. G. Lore, A. Akintayo, and S. Sarkar, "Llnet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognition*, vol. 61, pp. 650–662, 2017.

[12] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," *IEEE transactions on image processing*, vol. 30, pp. 2340–2349, 2021.

[13] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 1777–1786.

[14] X. Zhu, Y. Xu, H. Xu, and C. Chen, "Quaternion convolutional neural networks," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 631–647.

[15] J. Hou, Z. Zhu, J. Hou, H. Liu, H. Zeng, and H. Yuan, "Global structure-aware diffusion process for low-light image enhancement," *Advances in Neural Information Processing Systems*, vol. 36, pp. 79 734–79 747, 2023.

[16] D. Zhou, Z. Yang, and Y. Yang, "Pyramid diffusion models for low-light image enhancement," *arXiv preprint arXiv:2305.10028*, 2023.

[17] A. Niu, T. X. Pham, K. Zhang, J. Sun, Y. Zhu, Q. Yan, I. Kweon, and Y. Zhang, "Acdmsr: Accelerated conditional diffusion models for single image super-resolution," *IEEE Transactions on Broadcasting*, vol. 70, pp. 492–504, 2023.

[18] Z. Chen, Y. Zhang, D. Liu, J. Gu, L. Kong, X. Yuan *et al.*, "Hierarchical integration diffusion model for realistic image deblurring," *Advances in neural information processing systems*, vol. 36, 2024.

[19] Q. Yan, T. Hu, Y. Sun, H. Tang, Y. Zhu, W. Dong, L. V. Gool, and Y. Zhang, "Toward high-quality hdr deghosting with conditional diffusion models," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, pp. 4011–4026, 2023.

[20] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," *IEEE Transactions on image processing*, vol. 26, no. 2, pp. 982–993, 2016.

[21] S. Ma, W. Pan, N. Li, S. Du, H. Liu, B. Xu, C. Xu, and X. Li, "Low-light image enhancement using retinex-based network with attention mechanism." *International Journal of Advanced Computer Science & Applications*, vol. 15, no. 1, 2024.

[22] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, "Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 10 561–10 570.

[23] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *CVPR*, 2022.

[24] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 3060–3069.

[25] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," *IEEE Transactions on Image Processing*, vol. 30, pp. 2340–2349, 2021.

[26] S. Zhou, C. Li, and C. Change Loy, "Lednet: Joint low-light enhancement and deblurring in the dark," in *European conference on computer vision.* Springer, 2022, pp. 573–589.

[27] X. Xu, R. Wang, C.-W. Fu, and J. Jia, "Snr-aware low-light image enhancement," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 17 693–17 703.

[28] Y. Wang, R. Wan, W. Yang, H. Li, L.-P. Chau, and A. Kot, "Low-light image enhancement with normalizing flow," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 36, no. 3, 2022, pp. 2604–2612.

[29] T. Wang, K. Zhang, T. Shen, W. Luo, B. Stenger, and T. Lu, "Ultra-high-definition low-light image enhancement: A benchmark and transformer-based method," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 37, no. 3, 2023, pp. 2654–2662.

[30] Z. Fu, Y. Yang, X. Tu, Y. Huang, X. Ding, and K.-K. Ma, "Learning a simple low-light image enhancer from paired low-light instances," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 22 252–22 261.

[31] M. Liu, X. Li, and Y. Fang, "Low-light image enhancement via dual information-based networks," *Electronics*, vol. 13, no. 18, 2024. [Online]. Available: https://www.mdpi.com/2079-9292/13/18/3713

[32] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," *CoRR*, vol. abs/2001.06826, 2020. [Online]. Available: https://arxiv.org/abs/2001.06826

[33] X. Li, W. Wang, X. Feng, and M. Li, "Deep parametric retinex decomposition model for low-light image enhancement," *Computer Vision and Image Understanding*, vol. 241, p. 103948, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1077314224000298

[34] X. Wei, X. Zhang, S. Wang, C. Cheng, Y. Huang, K. Yang, and Y. Li, "Blnet: A fast deep learning framework for low-light image enhancement with noise removal and color restoration," 2021. [Online]. Available: https://arxiv.org/abs/2106.15953

[35] M. Zhu, P. Pan, W. Chen, and Y. Yang, "Eemefn: Low-light image enhancement via edge-enhanced multi-exposure fusion network," in *AAAI Conference on Artificial Intelligence*, 2020.

[36] Z. Zhu, X. Yang, R. Lu, T. Shen, T. Zhang, and S. Wang, "Ghost imaging in the dark: A multi-illumination estimation network for low-light image enhancement," *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[37] M. Y. Abbass, H. Kasban, and Z. F. Elsharkawy, "Low-light image enhancement via improved lightweight yuv attention network," *Computers & Graphics*, p. 104170, 2025.

[38] Y. Ma, S. Xie, W. Xu, X. Chen, X. Huang, Y. Sun, and W. Liu, "Region-based unsupervised low-light image enhancement in the wild with explicit domain supervision," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[39] H. Wang, X.-H. Wang, Y. Zhou, and J. Yang, "Color texture segmentation using quaternion-gabor filters," in *2006 International Conference on Image Processing*, 2006, pp. 745–748.

[40] S. J. Sangwine and T. A. Ell, "Hypercomplex fourier transforms of color images," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, vol. 1. IEEE, 2001, pp. 137–140.

[41] P. Fletcher and S. J. Sangwine, "The development of the quaternion wavelet transform," *Signal Processing*, vol. 136, pp. 2–15, 2017.

[42] Q. Yin, J. Wang, X. Luo, J. Zhai, S. K. Jha, and Y.-Q. Shi, "Quaternion convolutional neural network for color image classification and forensics," *IEEE Access*, vol. 7, pp. 20 293–20 301, 2019.

[43] W. R. Hamilton, "On quaternions; or on a new system of imaginaries in algebra," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 25, no. 163, pp. 10–13, 1844.

[44] Y. Feng, C. Zhang, P. Wang, P. Wu, Q. Yan, and Y. Zhang, "You only need one color space: An efficient network for low-light image enhancement," 2024.

[45] J. Qu, R. W. Liu, Y. Gao, Y. Guo, F. Zhu, and F.-Y. Wang, "Double domain guided real-time low-light image enhancement for ultra-high-definition transportation surveillance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 8, pp. 9550–9562, 2024.

[46] L. Kaiser, A. N. Gomez, and F. Chollet, "Depthwise separable convolutions for neural machine translation," *arXiv preprint arXiv:1706.03059*, 2017.

[47] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proceedings of the 27th ACM International Conference on Multimedia*, ser. MM '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1632–1640.

[48] H. Zeng, J. Cai, L. Li, Z. Cao, and L. Zhang, "Learning image-adaptive 3d lookup tables for high performance photo enhancement in real-time," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1–1, 2020. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2020.3026740

Appendix A

Proof of Publication

The following page(s) contain the proof of publication of the research work carried out as part of this thesis. The paper was published in the journal:

The acceptance letter / first page of the published article / certificate of publication is attached below.

Fwd: JVCI-25-32R1 - Final Decision

1 message

Prof. Anil Singh Parihar <parihar.anil@gmail.com>   Wed, May 7, 2025 at 5:06 PM
To: Asst Prof Kavinder Singh <kavinder@dtu.ac.in>, "sudeep05032@gmail.com" <sudeep05032@gmail.com>

Prof. ANIL SINGH PARIHAR
Professor
Department of Computer Science & Engineering
Associate Head, Dept. of Training & Placement
Delhi Technological University
(*formerly* Delhi College of Engineering)

---------- Forwarded message ---------
From: **JVIS** <em@editorialmanager.com>
Date: Tue, 6 May 2025 at 11:49 PM
Subject: JVCI-25-32R1 - Final Decision
To: ANIL SINGH PARIHAR <parihar.anil@gmail.com>

Ms. No.: **JVCI-25-32R1**
Title: QLight-Net: Quaternion based Low Light Image Enhancement Network
Corresponding Author: Dr. ANIL SINGH PARIHAR
Authors: Sudeep Kumar Acharjee; Dr. Kavinder Singh

Dear Dr. PARIHAR,

We are pleased to inform you that your manuscript referenced above has been accepted for publication in the Journal of Visual Communication and Image Representation.

Your accepted manuscript will now be transferred to our production department and work will begin on creation of the proof. If we need any additional information to create the proof, we will let you know. If not, you will be contacted again in the next few days with a request to approve the proof and to complete a number of online forms that are required for publication.

We appreciate and value your contribution to Journal of Visual Communication and Image Representation

Comments to Authors:

Reviewer's Responses to Questions

Note: In order to effectively convey your recommendations for improvement to the author(s), and help editors make well-informed and efficient decisions, we ask you to answer the following specific questions about the manuscript and provide additional suggestions where appropriate.

1. Are the objectives and the rationale of the study clearly stated?

Please provide suggestions to the author(s) on how to improve the clarity of the objectives and rationale of the study. Please number each suggestion so that author(s) can more easily respond.

Reviewer #1: 1. In section 2.1.1 of the manuscript, there are equations that are not numbered.

Reviewer #2: yes

Reviewer #3: na

Full Length Article

QLight-Net: Quaternion based low light image enhancement network☆

Sudeep Kumar Acharjee, Kavinder Singh [ID], Anil Singh Parihar [ID] *

Department of Computer Science and Engineering, Delhi Technological University, New Delhi, India

ARTICLE INFO

ABSTRACT

Images captured at night suffer from various degradations such as color distortion, low contrast, and noise. Many existing methods improve low-light images may sometimes amplify noise, cause color distortion, and lack finer details. The existing methods require larger number of parameters, which limits the adoption of these methods in vision-based applications. In this paper, we proposed a QLight-Net method to achieve a better enhancement with a comparably lower number of parameters. We proposed depth-wise quaternion convolution, and quaternion cross attention to develop the two-branch architecture for low-light image enhancement. The proposed model leverages gradient branch to extract color-aware gradient features. Further, It uses color branch to extract gradient-aware color features. The proposed method achieves an LPIPS score of 0.047, which surpasses the previous best results with lesser parameters, and achieves 0.88 and 29.05 scores of SSIM and PSNR, respectively. Our approach achieves a balance between computational efficiency and better enhancement.

1. Introduction

The rapid growth of digital imaging in diverse fields, including intelligent systems, has driven a need for image enhancement. From military surveillance to underwater photography, digital images are often captured in challenging low-light conditions, necessitating reliable enhancement methods. Photographs taken outdoor in low-light environment may suffer from low-light, lens flare, or extra noise. Images captured in low-light conditions may affect the decision-making capabilities of autonomous vehicles. Low-Light image enhancement focuses mainly on improving the brightness while maintaining the color consistency and reducing noise, and impact of color bias [1].

In the past decades, an extensive development in the field of low-light image enhancement proposed variety of approaches, from traditional methods in the pre-deep learning era to using of diffusion models in recent years. One of the earliest methods [2] that introduces histogram equalization based method that proposed iterative histogram modification that improves the contrast of digital images. Early works laid the groundwork for modern image enhancement techniques that use gamma correction for brightness and retinex theory for enhancing images that suffer from low-light environment. Another work [3] created a foundational ground by introducing the Retinex theory, explaining how human vision perceives color under varying lighting conditions by decomposing and image into reflectance and illumination components. Later, Edwin Land expands on the Retinex theory [4]

by emphasizing how the visual system maintains color consistency under different lighting conditions and how it can be applied to image enhancement. Since the introduction of Convolutional Neural Networks (CNN), many CNN-based methods for low-light image enhancement have been proposed, and laid the groundwork for using CNN's for low-light image enhancement task. In LLNet [5], it introduced one of the initial CNN-based models for low-light enhancement, using a deep auto-encoder to improve brightness and reduce noise in the low-light image. RetinexNet [4] combines the concept of retinex theory with CNN to decomposes image into reflectance and illumination. The work focus on decomposing and adjusting the illumination component. There are rapid development in unsupervised methods. Jiang et al. [6] used GAN that enhanced extremely low-light images without paired training data.

While these approaches achieved significant performance improvement, they also present certain limitations that motivate this research. Most existing deep learning-based methods rely on large, complex architectures with millions of parameters. Such models are computationally expensive, which limits their deployment on edge devices such as autonomous vehicles and mobile applications. Although some lightweight models are proposed [7], many of them compromise on enhancement quality or fail to retain structural and color details. Furthermore, current approaches typically operate in real-valued space, resulting in a higher number of parameters. These models also struggle to explicitly model the interaction between different types of features,

---

Appendix B

Plagiarism Verification

Title of the Thesis _____ **"QLight-Net: Quaternion based Low Light Image** _____
Enhancement Network"

Total Pages _____ **34** _____ Name of the Scholar __**"Sudeep Kumar Acharjee"**__

Supervisor(s)
(1) __**"Prof. Anil Singh Parihar"**__

Department of _____ **"Computer Science and Engineering"** _____

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: _____ **"Turnitin"** _____ Similarity Index: _____ **11%** _____ Total Word Count: _____ **13667** _____

The acceptance letter / first page of the published article / certificate of publication is attached below.

Date: _____ **May 2025** _____

Candidate's Signature   **Signature of Supervisor**
(Sudeep Kumar Acharjee)   Prof. Anil Singh Parihar
                                                            Computer Science & Engineering
                                                            Delhi Technological University

Sudeep

Final_Thesis.pdf

Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:98462794

Submission Date

May 30, 2025, 7:16 AM GMT+5:30

Download Date

May 30, 2025, 7:20 AM GMT+5:30

File Name

Final_Thesis.pdf

File Size

17.7 MB

34 Pages

13,667 Words

74,510 Characters

11% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- ▸ Bibliography
- ▸ Small Matches (less than 8 words)

Exclusions

- ▸ 1 Excluded Source

Match Groups

- 🟥 **119** Not Cited or Quoted 9%
  Matches with neither in-text citation nor quotation marks

- 💬 **22** Missing Quotations 2%
  Matches that are still very similar to source material

- ▤ **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation

- ⬡ **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

7%	🌐	Internet sources
8%	📖	Publications
6%	👤	Submitted works (Student Papers)

Integrity Flags

1 Integrity Flag for Review

🚩 **Hidden Text**
1 suspect characters on 1 page
Text is altered to blend into the white background of the document.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

119 Not Cited or Quoted 9%
Matches with neither in-text citation nor quotation marks

22 Missing Quotations 2%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

7% 🌐 Internet sources

8% 📖 Publications

6% 👤 Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Internet	
arxiv.org		**1%**

2	Internet	
export.arxiv.org		**<1%**

3	Publication	
C. St. J. A. NASH-WILLIAMS, D. J. WHITE. "Rearrangement of vector series. I", Math...		**<1%**

4	Internet	
scholar.afit.edu		**<1%**

5	Submitted works	
Heriot-Watt University on 2025-04-18		**<1%**

6	Internet	
ebin.pub		**<1%**

7	Internet	
www.readkong.com		**<1%**

8	Publication	
Yuntae Jeon, Dai Quoc Tran, Khoa Tran Dang Vo, Jaehyun Jeon, Minsoo Park, Seun...		**<1%**

9	Publication	
Zhou, Zheng. "Deep Feature Learning for Image Denoising and Style Transfer", U...		**<1%**

10	Publication	
Gonghe Xiong, Shan Gai, Bofan Nie, Feilong Chen, Chengli Sun. "Non-local feature...		**<1%**

Sudeep

Final_Thesis.pdf

Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:98462794

Submission Date

May 30, 2025, 7:16 AM GMT+5:30

Download Date

May 30, 2025, 7:22 AM GMT+5:30

File Name

Final_Thesis.pdf

File Size

17.7 MB

34 Pages

13,667 Words

74,510 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely AI generated as AI generated and AI paraphrased or likely AI generated and AI paraphrased writing as only AI generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.